# PhD THESIS

# STRUCTURAL ANALYSIS OF U3 SMALL NUCLEOLAR RNA OF *CHLAMYDOMONAS REINHARDTII IN VIVO*: FUNCTIONAL IMPLICATIONS

WRITTEN BY

MÁRIA ANTAL

SUPERVISORS:

FERENC SOLYMOSY

TAMÁS KISS

INSTITUTE OF PLANT BIOLOGY

BIOLOGICAL RESEARCH CENTER OF

HUNGARIAN ACADEMY OF SCIENCES

SZEGED, HUNGARY

2001

CONTENT

# 1 SUMMARY

U3 small nucleolar RNA (U3 snoRNA) is the most abundant stable RNA in the box C/D snoRNA family. In the form of ribonucleoprotein (RNP) complex it plays an essential role, as chaperone, in an ordered series of cleavages leading to the formation of mature 17/18S, 5.8S and 25/28S rRNAs in the nucleolus.

Little information is available on the nature of the association with the U3 snoRNA of either the common box C/D-associated proteins or the U3 specific proteins. In the current study, a detailed comparison of the chemical and enzymatic modification patterns obtained under *in vivo* conditions and, as a control, in solution of the U3 snoRNA from *Chlamydomonas reinhardtii* (*Cre*) allowed us: a) to identify the two-dimensional structure of and the potential protein binding sites in the *Cre* U3 snoRNA and b) to critically evaluate recently published models for direct base-pairing interactions between the 18S part of rRNA and U3 snoRNA.

**CONCLUSIONS:**

THE 5' DOMAIN OF *CRE* U3 snoRNA

/1/ Of all the U3 snoRNAs identified so far, *Cre* U3 snoRNA shows the strongest similarity to the higher plant U3 snoRNAs.

/2/ The 5' domain of plant U3 snoRNAs was proposed to be folded into a two stem-loop structure. Our studies revealed that the two hairpin structure of the 5' domain of the *Cre* U3 snoRNA can be confirmed in solution, but is no longer detected under *in vivo* conditions: the region between positions 53 and 67 is single-stranded.

/3/ Compared to earlier models for higher plant U3 snoRNAs, *Cre* U3 snoRNA shows extended hinges 1 and 2.

/4/ On the basis of our experimental results obtained under *in vivo* conditions a new modified model for the interaction of *Cre* U3 snoRNA with the 18S part of the pre-rRNA is proposed.

THE 3' DOMAIN OF *CRE* U3 snoRNA

/5/ Our structure probing data confirmed the existence of the predicted Stem/Loop II, the Central Stem and the 3' Stem structures. The Stem II region, although highly structured *in vitro,* appeared to be more relaxed *in vivo.*

/6/ In the 3' domain of *Cre* U3 snoRNA, only the 3' Stem and the box B regions are folded identically under *in vitro* and *in vivo* conditions.

/7/ The Stem/Loop III region exhibited variable reaction patterns in solution. Beside the WT-1 structure this part of the molecule tended to adopt an alternative two-dimensional structure called the Wild-Type 2 (WT-2) structure. We defined the structural elements responsible for the instability of Stem/Loop III by constructing and analyzing a series of substitution **(M)** and deletion **(MΔ)** mutants.

/8/ Under *in vivo* conditions the 3' domain of *Cre* U3 snoRNA folds into a two-dimensional structure that is highly reminiscent of the WT-1 structure. Most likely, snoRNP proteins bound to the U3 RNA inhibit the formation of the alternative WT-2 structure and stabilize the WT-1 (or a WT-1-like) structural conformation in the nucleoli of *Cre* cells.

/9/ The box C'/D and B/C motifs play an essential role in protein binding. Their accessibility shows significant differences in the presence of proteins as compared to their behavior *in vitro. In vivo* both motifs can adopt the structural characteristics of a 15.5kD protein binding site. Owing to the modification pattern of adenines and cytosines the direct protein binding site could not be defined at nucleotide level.

/10/ In solution, the box C sequence of *Cre* U3 snoRNA, if in the WT-1 conformation, is able to form a five base-pair interaction with the Loop III region, resulting in pseudoknot formation. After a systematic inspection, we found that, in principle, all known U3 snoRNAs could form a box C - Loop III-type pseudoknot structure which could function as an excellent protein docking signal.

## 2 INTRODUCTION

RNA molecules are known to have wide range of functions to in the cells. While the ribosomal RNAs (rRNAs) constitute approximately 80% of the total RNA in rapidly growing mammalian cells (e.g. cultured HeLa cells), the transfer RNA (tRNA) content is 15% and protein-coding messenger RNAs (mRNAs) give only a small portion. The small nuclear and nucleolar RNAs (snRNAs, snoRNAs) represent even less abundant groups of stable RNAs. **SnRNAs** are localized in the nucleoplasm and play an essential role in pre-mRNA splicing, when the intron sequences are excised by two sequential transesterification reactions (Staley and Guthrie, 1998). **SnoRNAs** are localized in the nucleolus, where the mature 18S, 5.8S and 28S ribosomal RNAs are synthesized from precursor rRNA (pre-rRNA) via a series of complex processing events, reviewed in (Eichler and Craig, 1994; Raue and Planta, 1995)(see Fig. 1).

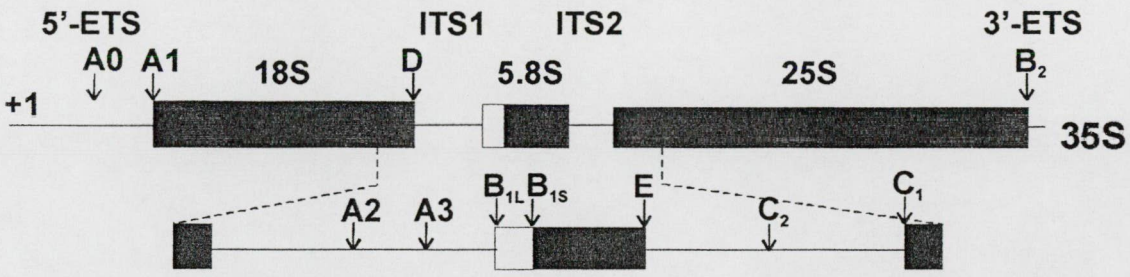


**Figure 1.** Schematic representation of the processing pathway of 35S pre-rRNA in *Saccharomyces cerevisiae*. The mature 18S, 5.8S and 25S rRNA sequences are shown as thick bars, external and internal transcribed spacer sequences (5'-ETS, 3'-ETS, ITS1 and ITS2) as thin bars. Cleavage sites are indicated by arrows and by uppercase letters (Kressler et al., 1999; Venema and Tollervey, 1999).

The majority of snoRNAs function as sequence-specific guides in the modification of rRNA and only few of them are essential for pre-rRNA cleavage events. Based on conserved sequence/secondary structure elements, more than 150 snoRNAs (with the exception of MRP RNA) belong to either the box H/ACA or the C/D snoRNA family and both function in the form of small nucleolar ribonucleoprotein particles (snoRNP) (Lafontaine and Tollervey, 2000; Watkins et al., 1998a). While the **box H/ACA snoRNAs** direct the pseudouridylation (Bortolin et al., 1999; Ganot et al.,

1997a; Ganot et al., 1997b), the **box C/D snoRNAs** guide the 2′-O-methylation of the rRNA (Kiss-Laszlo et al., 1998; Kiss-Laszlo et al., 1996). The mature rRNAs carry about 110 2'-O-methyl groups and almost a hundred pseudouridines (Maden, 1990).

**U3 snoRNA**, the subject of the present thesis, is a member of the **box C/D** family. The **U3 snoRNA gene** and the telomerase RNA gene are the only known examples of genes that are transcribed by different RNA polymerases in different organisms (for the telomerase see Greider, 1996; Chapon et al., 1997). In vertebrates and yeasts U3 snoRNA is transcribed by RNA polymerase II (Busch et al., 1982; Hughes et al., 1987) and in higher plants by RNA polymerase III (Filipowicz et al., 1990; Kiss et al., 1991). This difference is also reflected by the presence of two different 5'-cap structures: $m^{2,2,7}$Gppp (TMG)- in vertebrates and yeasts and γmppp- in plants (Busch et al., 1982; Myslinski et al., 1990; Shimba et al., 1992). It was, thus, interesting to know at what stage of evolution the change in RNA polymerase specificity took place. In our laboratory we chose the *Chlamydomonas reinhardtii* U3 snoRNA (*Cre* U3 snoRNA) as a model object to test the enzyme specificity of U3 RNA gene transcription. The study of a U3 gene of a unicellular alga, probably a remote ancestor of higher plants, was expected to shed some light on this problem. Combining direct U3 snoRNA 3'-terminal sequence analysis, cDNA synthesis and sequence analysis, the complete sequence of *Cre* U3 snoRNA was established. The cDNA generated with U3 snoRNA as a template was used to isolate a U3 gene from a *Cre* genomic library. *Cre* cell transformation experiments indicated that the isolated gene was an expressed *bona fide* U3 gene. The results obtained strongly support the idea that the *Cre* U3 gene is a **pol III transcript** (Antal et al., 2000).

The U3 snoRNA functions in the form of **U3 snoRNP** *in vivo* (Tyc and Steitz, 1989; Kiss and Solymosy, 1990). The RNA is associated with **proteins common** to all box C/D snoRNAs, including fibrillarin/Nop1p (Schimmang et al., 1989; Baserga et al., 1991), Nop56 (Lafontaine and Tollervey, 2000), and Nop58/Nop5p (Lafontaine and Tollervey, 1999; Lyman et al., 1999; Newman et al., 2001) and **proteins specific** to the U3 snoRNP, including Sof1p (suppressor of fibrillarin (Jansen et al., 1993)), Mpp10 (Wormsley et al., 2001), Lcp5p (Wiederkehr et al., 1998), Imp3p (Lee and Baserga, 1999), Imp4p (Lee and Baserga, 1999), Dhr1p (Colley et al., 2001) and U3-55k/Rrp9p (Lübben et al., 1993; Pluk et al., 1998; Venema et al., 2001). Each of these proteins

(like U3 RNA itself) is essential and has been shown to be specifically required for maturation of pre-rRNA (Tollervey et al., 1991; Gautier et al., 1997).

Among the snoRNAs functionally involved in the ordered series of cleavages leading to mature 17/18S, 5.8S and 25/28S rRNAs, U3 snoRNA plays an essential role (see Fig. 1). Abolition of U3 snoRNA gene expression in *S. cerevisiae* blocks 18S rRNA synthesis, whereas the 5.8S and 25S rRNAs are produced in normal amounts (Hughes and Ares, Jr. 1991). U3 snoRNA is involved in the very early cleavage (A0), that takes place in the 5'-external transcribed spacer (5'-ETS) of pre-rRNA, as demonstrated in mouse cell extracts, in *Xenopus* oocyte extracts and in yeast (Craig et al., 1987; Kass et al., 1990; Mougey et al., 1993; Beltrame and Tollervey, 1995). U3 snoRNA is also involved in cleavages occurring on both sides of the 18S rRNA coding region (A1 and A2) as shown in yeast and in *Xenopus* oocytes (Hughes and Ares, Jr. 1991; Savino and Gerbi, 1990) and at a site near the boundary of the internal transcribed spacer 1 (ITS1) and 5.8S rRNA, as found in *Xenopus* oocytes (Savino and Gerbi, 1990). Based on *in vivo* crosslinking experiments, a base-pair interaction was first discovered between U3 snoRNA and the 5'-ETS region (Fig. 2/A – helix V)(Beltrame and Tollervey, 1992).

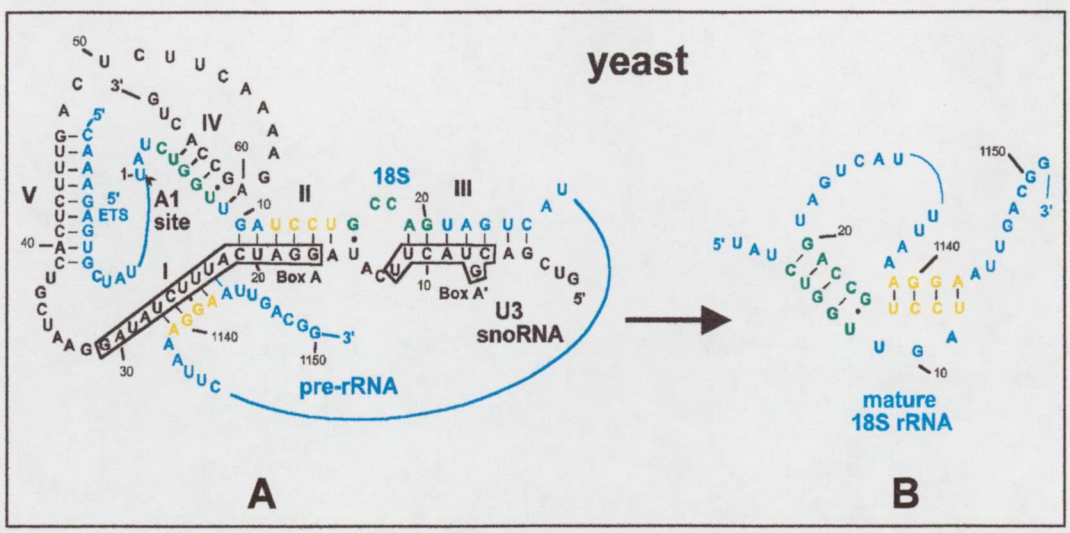


**Figure 2. (A)** Proposed interaction between the *Saccharomyces cerevisiae* U3 snoRNA (in black) and pre-rRNA (in color). The interaction is via four base-paired helices between the U3 and the 18SrRNA (helices I, II, III (Hughes, 1996) and IV (Méreau et al., 1997)) and one base-paired helix between the U3 and the 5'-ETS of pre-rRNA (Beltrame et al., 1994). Boxes A and A' are framed. **(B)** Formation of a pseudoknot by the partner sequences in 18S rRNA as proposed by Gutell et al., (Gutell et al., 1994).

The biological significance of this interaction was largely documented by genetic experiments, phylogenetic evidence and direct *in vivo* analysis of the U3 snoRNA structure in yeast (Beltrame and Tollervey, 1995; Brule et al., 1996; Méreau et al., 1997). In a phylogenetic study, three intermolecular base-pair interactions (helices I, II and III) between U3 snoRNA and pre-rRNA were proposed (Hughes, 1996) in which the U3 snoRNA segments involved are contained in the phylogenetically conserved boxes A and A' and the pre-rRNA segments are located in the 18S rRNA region (Fig. 2/A). These segments form a pseudoknot structure in mature 18S rRNA (Fig. 2/B)(Hughes, 1996). Later on, on the basis of an *in vivo* analysis of the *S. cerevisiae* U3 snoRNA structure, the model proposed by Hughes was revised (helices I, II and III were extended) and a fourth bimolecular interaction (helix IV) that involves the fourth segment of the 18S rRNA pseudoknot structure was proposed (Méreau et al., 1997). Mutation in the U3 snoRNA region involved in the formation of helices I, II and III (Hughes, 1996; Méreau et al., 1997; Samarsky and Fournier, 1998) demonstrated the functional importance of these U3 snoRNA sequences. More recently, the importance of helix II was demonstrated by the generation of compensatory mutations in the U3 snoRNA and pre-rRNA segments involved in its formation (Sharma and Tollervey, 1999). When the same kind of experiment was performed for helix I, no compensatory effect was observed (Sharma and Tollervey, 1999).

**U3 snoRNA** shows a high degree of **structural conservation** during evolution. All U3 snoRNAs characterized so far carry six evolutionarily conserved sequence elements. In addition to the U3-specific boxes (A, A' and B sequence motifs), the conserved boxes C, C' and D are common to the box C/D snoRNA family (Kiss and Solymosy, 1990; Hartshorne and Agabian, 1994; Samarsky and Fournier, 1998; Maxwell and Fournier, 1995). The two-dimensional structure of all U3 snoRNAs can be divided into two (5' and 3') major domains (Parker and Steitz, 1987; Jeppesen et al., 1988; Segault et al., 1992; Hartshorne and Agabian, 1994; Méreau et al., 1997; Fournier et al., 1998).

The **5' domain** contains the boxes A' and A which are capable of forming direct base-pairing interactions with 18S rRNA sequences facilitating the correct folding of a pseudoknot structure in the 18S rRNA as described above (see Fig. 2)(Hughes, 1996; Méreau et al., 1997; Parker and Steitz, 1987; Beltrame et al., 1994; Beltrame and Tollervey, 1995; Antal et al., 2000). Another non-conserved single-stranded stretch in

the 5' domain of the U3 snoRNA is believed to base-pair with a short fragment of the 5'-ETS (external transcribed spacer) of pre-rRNA (Beltrame and Tollervey, 1995; Dennis et al., 1997; Hartshorne, 1998; Intine et al., 1999; Borovjagin and Gerbi, 1999) and, therefore, to contribute to the first endonucleolytic cleavage of the primary rRNA transcript. The secondary structure of the 5' domain of U3 snoRNAs is not uniform in different organisms; it can be folded into a single stem/loop structure in higher eukaryotes or a two stem/loop structure in yeast and plants. In protists, this region of U3 shows no apparent secondary structure. The 5' domain of U3 snoRNA is required for the function but not for the accumulation of the RNA (Samarsky and Fournier, 1998). A single-stranded hinge region follows the 5' domain of U3 snoRNA, which is one of the least understood segments of the U3 snoRNA. It most likely provides the correct spacing between the 5' and 3' domains of the RNA (Borovjagin and Gerbi, 2000; Antal et al., 2000). Although the primary sequence of the hinge region differs among eukaryotic organisms, its lengths is fairly well conserved (between 34 and 37nt) (Samarsky and Fournier, 1998). Recently this region was found to be required for association with protein Mpp10p (Wormsley et al., 2001).

In the highly structured **3' domain**, the conserved boxes B, C, C' and D are believed to function as snoRNP protein anchoring signals (Hartshorne and Agabian, 1994, Jeppesen et al., 1988; Méreau et al., 1997; Parker and Steitz, 1987; Lübben et al., 1993). The secondary structure of the 3' domain positions boxes B and C, as well as boxes C' and D into discrete loops flanked by base-paired stems to form the box B/C and box C'/D motifs, respectively (Méreau et al., 1997; Samarsky and Fournier, 1998; Speckmann et al., 1999). The box C'/D motif of U3 is equivalent to the box C/D motif of other members of this snoRNA family in sequence/structure and function (Méreau et al., 1997; Samarsky and Fournier, 1998; Speckmann et al., 1999). The box B/C motif is unique to U3 snoRNA. Both the box B/C and box C'/D motifs operate as independent units of U3 RNA in nuclear retention in Xenopus oocyte (Speckmann et al., 1999). The box C'/D motif of U3 also directs the localization of the snoRNA to the nucleolus (Narayanan et al., 1999). The functions associated with the box B/C (Venema et al., 2001; Lukowiak et al., 2000; Colley et al., 2001) and box C'/D sequence motifs are very likely mediated by proteins that bind at these sites. In conjunction with the flanking 3'-terminal and central stem structures, boxes C' and D are critical for U3 accumulation (Samarsky and Fournier, 1998). Alterations introduced into the box D sequences and

destruction of the 3'-terminal stem structure block TMG cap formation in *Xenopus* U3 snoRNA and impair its nuclear retention (Terns and Dahlberg, 1994). The presence of a correct box C motif is essential for fibrillarin binding (Baserga et al., 1991). Neither the number - from none in protists to three in yeast - nor the sequence of the surrounding hairpins is conserved and none is critical for the accumulation or the function of the U3 snoRNA (Fournier et al., 1998; Samarsky and Fournier, 1998).

Recently, several models implicating direct snoRNA-rRNA interactions have been proposed for the function of U3 snoRNA and more information has become available concerning U3 regions involved in RNA-protein interactions in the U3 snoRNP particle. However, the exact primary and secondary structural requirements for snoRNP formation or U3 snoRNA function have not yet been elucidated. In the current study, the structure of the 5' and 3' domains of U3 snoRNA from *Chlamydomonas reinhardtii* has been analyzed under *in vivo* conditions and as a control in solution in order to check the validity and phylogenetic impact of recent models. Our work provides for the first time experimental data on the higher order structure of a plant U3 snoRNA with special reference to probable protein-binding sites.

# 3 MATERIALS AND METHODS

## 3.1 Cell strains and culture conditions

Vegetative cells of *Chlamydomonas reinhardtii* CW15 strain (a kind gift from Dr CF Beck) were grown at 25 °C under continuous light in Tris-acetate-phosphate (TAP) medium on a gyratory shaker (Rochaix et al., 1988).

## 3.2 Isolation of U3 snoRNA from *Chlamydomonas reinhardtii (Cre)* cells and nucleotide sequence determination

The cells from 2L of culture in log phase (2 x $10^6$ cells/ml) were collected by centrifugation (5000g for 10 min). Total RNA was isolated by the acidic guanidium-HCl-phenol/chloroform method (Rochaix et al., 1988) with two modifications: The high molecular weight RNAs were removed by precipitation with 20% ethanol, and the low molecular weight RNAs left in the supernatant fraction were precipitated by adding 3 volume of 96% ethanol instead of 2M LiCl. The low molecular weight RNAs were separated on 10% polyacrylamide-8M urea gels and the RNAs were detected by staining with ethidium bromide. The band corresponding to U3 snoRNA, was cut out from the gel. Elution from the gel slices, 3'-end labeling and chemical sequencing of the 3' portion of the RNA were done as described in Kiss et al. (1987). For the synthesis of a complete cDNA corresponding to the *Cre* U3 snoRNA, the first strand was obtained with the synthetic oligonucleotide primer (P/3'Eco) complementary to the 3' end of the RNA. Elongation with M-MLV Reverse Transcriptase (Gibco BRL) was done according to the manufacturer's protocol, with addition of 10 μCi of [α-$^{32}$P]dCTP. After completion of the reaction, 1/20 vol. of 3 M NaOH was added to the mixture, followed by boiling for 5 minutes. The labeled first-strand cDNA was repurified from a 10% polyacrylamide-8 M urea gel. The recovered single-stranded cDNA was dG tailed using terminal-deoxynucleotidyl transferase (USB) according to Eschenfeldt *et al.* (1987), and PCR amplified by Vent DNA polymerase (Biolabs) using a synthetic oligonucleotide [oligo $(dC)_{12}$/*Eco*RI] as the second primer. After purification on a 2% agarose gel, the amplified DNA fragment was cloned into the *Sma*I site of the pBS (-) vector **pCRU3C** (Stratagene). For double-stranded DNA sequencing Sequenase (USB) was used according to the supplier. Identification of the cap structure carried by the *Cre*

U3 snoRNA was achieved by immune-affinity chromatography using anti-$m^{2,2,7}G$ IgG and anti-γmpppN IgG, as previously described in (Jakab et al., 1997; Liu et al., 1992), respectively.

### 3.3 Secondary structure alignment of the 5' domain of U3 snoRNA from *Chlamydomonas reinhardtii* and from other organisms

Different U3 snoRNA sequences proposed to be folded into a two stem-loop structure were in a first step treated with the program described by Lück *et al.* (1999), then, alignment was refined manually in order to align the phylogenetically conserved boxes and for a better representation of compensatory mutations.

### 3.4 Plasmid constructions for *in vitro* transcriptions

To get an *in vitro* transcript of the wild-type (wt) *Cre* U3 snoRNA, we subjected the pCRU3C plasmid to PCR amplification using the 5' homologous oligonucleotide (P/Bam/T7/5') and oligonucleotide P/3'Eco. The PCR fragment was cloned into the BamHI/EcoRI sites of pBSKII(+) (**pCRU3WT**). The plasmid was transcribed by T7 RNA polymerase after linearization with EcoRI. Owing to the location of the EcoRI restriction site, the U3 snoRNA transcript contained five additional nucleotides at its 3' end. The pCRU3WT plasmid served as a template for all *in vitro* mutagenesis experiments (Kunkel et al, 1987) using the oligonucleotides listed in Table 2.

### 3.5 *In vitro* transcription of U3 snoRNAs

*In vitro* transcription was carried out in 80 mM Tris-HCl (pH 7.9), 20 mM DTT, 20 mM NaCl, 2 mM spermidine, 12 mM $MgCl_2$, on 2 μg of linearized plasmid, with 20 units of RNasin, 1.5 mM rNTPs (Boehringer) and 69 units of T7 RNA polymerase (NEB) in 20 μl final volume. After one hour incubation at 37 °C the DNA template was digested with five units of RNase free DNase I (Boehringer) in 200 μl of 5 mM $MgCl_2$, 100 mM sodium acetate (pH 5.0) buffer for 30 minutes at 37 °C. The RNA transcripts were phenol-extracted after adding 0.2% SDS, ethanol precipitated and dissolved in sterile water (Chabot, 1994).

**3.6 Structure probing *in vitro*. A) Chemical modifications and enzymatic digestions of unlabeled RNA transcripts**

30-80 ng of *in vitro* transcripts were used for each chemical modification or enzymatic digestion. This aliquot was diluted in the appropriate buffer without $Mg^{2+}$, denatured at 90 °C for 30 sec, 65 °C for one minute and cooled down to room temperature by addition of 10 mM $MgCl_2$. For each reaction, a control (no reagent) was treated in parallel to detect nicks in the RNA and pauses of reverse transcription.

**DMS** (dimethylsulfate from Aldrich) modification was carried out in 150 µl of 50 mM sodium cacodylate (pH 7.5), 10 mM $MgCl_2$ , 50 mM KCl, by the addition of 10 µg yeast tRNA (Boehringer) and 0.5 µl or 2 µl of DMS solution [1/1 (v/v) (DMS/EtOH)] at 20 °C for 15 minutes. The reaction was stopped with 60 µl of G-stop buffer [0.5 M 2-β-mercaptoethanol, 0.75 M sodium acetate (pH 5.5)] and ethanol-precipitated.

**Kethoxal** (2-keto-3-ethoxybutylaldehyde from USB) modification was performed in 80 µl of DMS buffer by the addition of 1.25 µg tRNA and 10 µl or 20 µl of 37 mg/ml freshly prepared kethoxal solution at 20 °C for 6 minutes. The reaction was stopped with 32 µl of 500 mM potassium borate (pH 7.0), 10 µg tRNA. After ethanol-precipitation the RNA was dissolved in 25 mM potassium borate (pH 7.0).

**CMCT** (N-cyclohexyl-N'-[b-(N-methylmorpholyno)-ethyl]-carbodiimide-p-toluenesulfonate) from Merck ) modification was carried out in 300 µl of 50 mM sodium borate (pH 8.0), 10 mM $MgCl_2$ , 50 mM KCl by the addition of 15 µg tRNA and 20 µl or 30 µl of 168 mg/ml CMCT solution (in reaction buffer) at 20 °C for 30 minutes. The RNA was ethanol-precipitated from 0.3 M sodium acetate.

Partial enzymatic digestion with **RNase V1** (Cobra venom V1 from Pharmacia) was carried out in 50 µl of 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$ , 50 mM KCl in the presence of 2.5 µg tRNA and 0.5 unit or 1.5units of RNase V1 at 20 °C for 10 minutes. The reaction was stopped with 3 µl of 100 mM EDTA, 10 µg tRNA . After phenol extraction the RNA was ethanol- precipitated.

**Lead cleavages** were induced on 0.1 µg of RNA transcript in 20 µl of 50 mM Hepes-KOH (pH 7.5), 10 mM magnesium acetate, 50 mM potassium acetate buffer in the presence of 10 µg tRNA at 25 °C for 15 minutes. Final concentrations of freshly-prepared lead acetate solution were 5 mM or 20 mM. The reaction was terminated by ethanol-precipitation in the presence of 33 mM EDTA (Ciesiolka, 1998).

**3.7 Structure probing *in vitro*. B) 3' and 5' end-labeling of RNA transcripts, partial enzymatic digestions and nucleotide sequence analysis**

10-15pmoles of gel-purified RNA transcripts were 3' end-labeled with 3', 5'-[5'-$^{32}$P] cytidine biphosphate (3000 Ci/mmole, pCp from Amerscham) by T4 RNA ligase (NEB) according to England et al. (1980). 10-15pmoles of dephosphorylated RNA transcripts were 5'end-labeled in the presence of 100 μCi of [γ-$^{32}$P] ATP (Dupont/NEN) and 10 units of T4 polynucleotide kinase (NEB) (Efstratiadis et al., 1977). After 3' and 5' end-labeling, the RNA samples were purified on 10% polyacrylamide-8 M urea gel. Structure probing of 50000 cpm 3' or 5' end-labeled RNAs in the presence of 2.5 μg tRNA was performed by limited digestions with 0.001-0.04units/μg of T1 (G specific), T2 (single-strand specific), and V1 (double-strand specific) RNases in 5μl of 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 50 mM KCl buffer at 20 °C for 10 minutes. The samples were directly loaded on 7% , 10% or 15% sequencing gels. As a control, limited alkaline hydrolysis and T1 RNA sequencing were performed according to Kiss et al. (1987).

**3.8 Structure probing *in vivo***

The procedure described by Zaug and Cech (1995) was used for structure probing analysis of *Cre* U3snoRNA *in vivo*. *Cre* cells were grown upto a concentration of $10^5$ cells/ml, harvested and resuspended at $10^8$ cells/ml concentration in 10 ml TMS buffer (10 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 3 mM $CaCl_2$, 250 mM sucrose). Two ml of this culture were treated with 20 μl, 40 μl, 80 μl of DMS, respectively, for 2 minutes at room temperature with gentle shaking. An aliquot treated in the same way, except addition of DMS, served as a control. 200 μl of 14.3 M β-mercaptoethanol (0.7 M final concentration) was added to quench the DMS. RNA was isolated from *Cre* cells as described above. The RNA pellet was dissolved at 10 mg/ml concentration and 1 μl was used for primer extension.

**3.9 Primer extension analysis**

20 ng of gel-purified oligonucleotide primer (Table 3), complementary to U3 snoRNA was 5' end-labeled at 37 °C for 30 minutes with 1 μl of [γ-$^{32}$P] ATP (3000 Ci/mmole)

by 10 units of T4 polynucleotide kinase. 0.5 ng of 5' end-labeled primer was annealed to chemically modified or partially digested RNAs at 90 °C for 30 sec, 65 °C for one minute and extended in 5 µl final volume in the presence of one unit of AMV reverse transcriptase (Life Science), 100 µM dNTPs, 50 mMTris-HCl (pH 8.3), 60 mM NaCl, 10 mM DTT, 6 mM $MgCl_2$ at 42 °C for 30 minutes. Positions of chemical modifications and enzymatic cleavages were identified by separation on 7% sequencing-gels (Biorad). Dideoxy sequencing ladder (50 µM ddNTPs) made on unmodified RNA with the same 5' end-labeled primer served as reference.

**Table 1.** Oligonucleotides used for plasmid constructions

| Denomination | Sequence of primer |
|---|---|
| P/3'Eco | 5' GGAATTCACCCATCTGAATGCCC 3' |
| P/*Bam*/T7/5' | 5'CGGGATCC**TAATACGACTCACTATAGG***ATGACCGT* |
| P/$(dC)_{12}$/EcoRI | 5' GGAATTCCCCCCCCCCCC 3' |

**Table 2.** Oligonucleotides used for Kunkel mutagenezis

| Denomination | Sequence of primer | Targeted sequence of U3 |
|---|---|---|
| P/ΔII | 5' GGT CTC GGC GGG \ GGC GCC GGG G 3' | 96 to 105 and 170 to 185 |
| P/StemIII 3' | 5' GCA CCA TCC CTA AAG GTG GAA TGG 3' | 180 to 203 |
| P/StemIII 3' + 5' | 5' CAC CAT CCC TAA AGG TGG AAT GGT CTC TTT GGG ACG TTC 3' | 164 to 202 |
| P/BoxB | 5' CAG TGG TCA TAC GTT TTG CTC AGG CGC CGG GGC 3' | 95 to 127 |
| P/BoxC | 5' GGG ACG TTC ATA AAG AAC CAC GCG 3' | 149 to 172 |

**Table 3.** Oligonucleotides used for primer extension

| Denominati on | Sequence of primer | Complementary to position of U3 |
|---|---|---|
| P/3' | 5' ACC CAT CTG AAT GCC CCG AGC | 199 to 222 |
| P/BoxC | 5' CTC GGC GGG ACG TTC 3' | 164 to 178 |
| P/BoxB | 5' CCA CAG TGG TCA TAC G 3' | 125 to 130 |
| P/Internal Loop | 5' GCA TGT CTT CGT CTG GGT TGA GGA 3' | 70 to 87 |

# 4 RESULTS

In eukaryotic cells different functions of RNA molecules require well defined RNA structures, coordinated RNA-RNA and RNA-protein interactions. The structure and function of an RNA can be studied by theoretical methods such as **phylogenetic comparative analysis**. This approach is based on sequence comparison of homologous RNAs from different organisms covering a large range through the evolution. It is a general method for highlighting conserved regions and gives the thermodynamically most favorable secondary structure models. Experimental procedures such as a) mutation-deletion analysis, b) X-ray crystallography, c) NMR and d) *in vivo* structure probing can support the theoretical implications. Physical approaches, such as X-ray and NMR require large amounts of highly purified material.

The **structure probing analysis**, a powerful method developed in the '80s for *in vitro* (Ehresmann et al., 1987) and *in vivo* conditions (Zaug and Cech, 1995; Inoue and Cech, 1985), can be applied to study the secondary and tertiary structure of an RNA molecule either free or engaged in complex with proteins. This unique method gives a detailed insight into the folding of an RNA by testing the reactivity of every nucleotide towards chemical or enzymatic probes. The detection of modified nucleotides and RNase cleavages can be conducted by two different paths. The first one uses end-labeled RNA molecule and allows detecting only scissions in the RNA chain. The second approach is based on primer extension by reverse transcriptase and detects stops of transcription at modified or cleaved nucleotides. The synthesized cDNA fragments are then separated by electrophoresis on sequencing polyacrylamide gels.

To obtain structural information on the *Cre* U3 snoRNA under *in vivo* conditions we modified intact CW15 algal cells with DMS. As a control we treated *in vitro* transcribed *Cre* U3 snoRNA - either 5' or 3' end-labeled and non-labeled- with chemical agents: DMS (N1 of adenine>>N3 of cytosine), CMCT (N3 of uracil>>N1 of guanine), kethoxal (N1 and N2 of guanine) (see Fig. 3), and with enzymes: RNase T1 (G) and RNase T2 or S1 nuclease (non-specific residues) as single strand-specific reagents (Ehresmann et al., 1987). Residues involved in base-pairing or stacking interactions were identified by partial RNase V1 (a double-strand-specific enzyme) digestions. The mapping of N7 position of guanine after DMS treatment gives information on the participation of this position in tertiary structure. The reactive

nucleotides in non-labeled U3 RNA samples were visualized by primer extension of end-labeled specific oligonucleotides, while the degradation products of terminally labeled U3 RNAs were directly analyzed by gel electrophoresis (data not shown). The occasional strong bands, found both *in vitro* and *in vivo,* represent nicked molecules. These bands can be considered as natural stops or pause of reverse transcriptase. The extra bands present only *in vivo* both in the control and in the DMS reaction are possibly generated by β-mercapto-ethanol used to quench the DMS (Zaug and Cech, 1995).

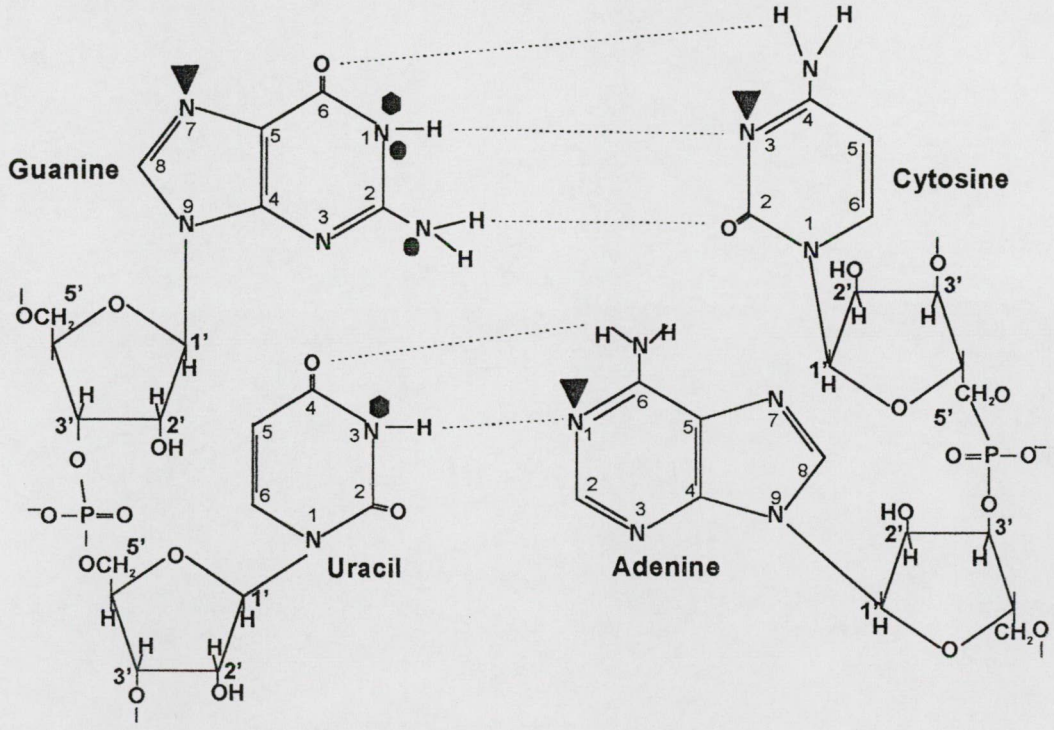


**Figure 3.** Canonical Watson-Crick interactions and target positions of the chemical probes. DMS (▼); CMCT (⬢); Kethoxal (●) (Ehresmann et al., 1987).

The structure probing analyses of the structurally and functionally distinct 5' and 3' domains of *Cre* U3 snoRNA are discussed below separately.

## 4.1 THE 5' DOMAIN OF *CRE* U3 SNORNA

### 4.1.1 Comparative sequence analysis of the 5' domain of *Cre* U3 snoRNA

In our studies we used a recently published program (ConStruct) developed by Lück et al. (Lück et al., 1999). This program yielded secondary structure models both for the 5' and for the 3' domain of *Cre* U3 snoRNA. The 222-nt long sequence of *Cre* U3 snoRNA was compared to those of U3 snoRNAs from other species. An alignment of the 5' portion of U3 snoRNAs according to secondary structure is shown in Fig. 4. The 75-nt long 5' domain of the *Cre* U3 snoRNA (denoted Crj001179) which contains the conserved boxes A' and A is assumed to be involved in base-pair interaction with phylogenetically conserved regions of the 18S moiety of pre-rRNA (Hughes, 1996; Méreau et al., 1997).

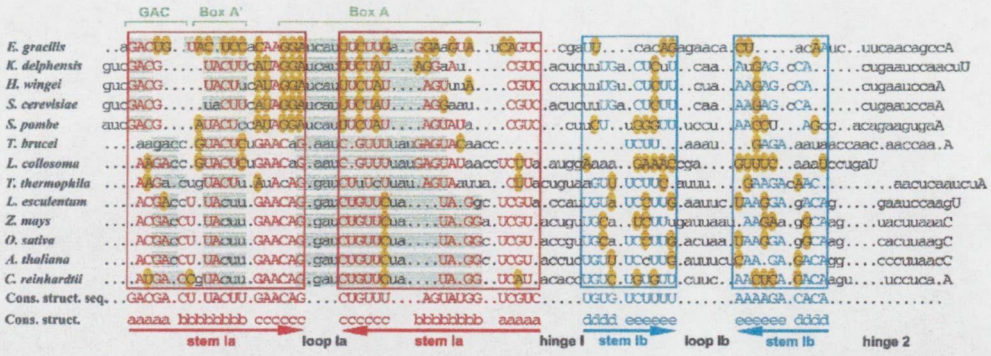


**Figure 4.** Compensatory base mutations conserved the possibility to form a two stem-loop structure in yeasts, ciliates, unicellular algae and plants U3 snoRNAs. The sequences of the 5'-terminal domain from *Euglena gracilis* (Eg27297), *Kluyveromyces delphensis* (Kdz78433), *Hansenula wingei* (Hwu3sno), *Saccharomyces cerevisiae* (Scsnr17a), *Schizosaccharomyces pombe* (Spsnru3), *Trypanosoma brucei* (Tburb), *Leptomonas collosoma* (Lctgv3snr), *Tetrahymena thermophila* (Ttsnru31), *Lycopersicon esculentum* (Leu3snr), *Zea mays* Zmu3snrng), *Oryza sativa* (Osu3snrn), *Arabidopsis thaliana* (Atu3csnr) and *Chlamydomonas reinhardtii* (Crj001179), were first subjected to the program developed by Lück et al. (1999). Then alignment was manually refined for a better alignment of the phylogenetically conserved boxes GAC, A' and A (highlighted in green) and for a better representation of covariations. The nucleotide sequences involved in the stem of stem-loop Ia are boxed in red. Stems are marked below by inverted arrows. The base-paired residues are in capital letters. Red capital letters correspond to residues fitting the consensus sequence established from the alignment . Nucleotide variations in this consensus sequence that were found to preserve base-pair interactions are circled in yellow. Dots indicate missing nucleotides. The same representation is used for stem-loop Ib, except that blue color is used instead of red color. The sequences can be accessed via the EMBL-ID at http://srs.ebi.ac.uk:5000/ site.

The high degree of sequence similarity observed between the 5' domain of *Cre* U3 snoRNA and its counterparts in U3 snoRNAs from other species is in line with this proposal. It should be noted that, whereas the GC content of most spliceosomal UsnRNAs from *Cre* is around 65 per cent, that of the 5' domain of U3 snoRNA is as low as 42.7 per cent. This is in support of the functional difference between the two domains of *Cre* U3 snoRNA. It can also be seen that with respect to both primary and proposed secondary structure, among all the U3 snoRNAs compared in Fig. 4, *Cre* U3 snoRNA shows the strongest similarity with the higher plant U3 snoRNAs. Whereas, the 5' domain of vertebrate U3 snoRNAs can be folded into a single stem-loop structure (not shown), the 5' domain of plant U3 snoRNAs was proposed to be folded into a two stem-loop structure (Fig. 4) (Solymosy and Pollák, 1993). The comparison shown in Fig. 4 predicts a two stem-loop structure for the 5' domain of *Cre* U3 snoRNA. In the present study, we give experimental evidence for this proposal.

### 4.1.2 Structure of the 5'-terminal domain of *Cre* U3 snoRNA *in vivo*

Recent literature data show that owing to its interaction with pre-rRNA, the 5'-terminal structure of yeast U3 snoRNA is different in semi-purified U3 snoRNP and *in vivo,* as well (Méreau et al., 1997). To see whether or not this is true of *Cre* U3 snoRNA, the structure of the 5'-terminal domain of *Cre* U3 snoRNA was probed i) in solution using chemical reagents and enzymes, and ii) *in vivo* using DMS. The *in vitro* study of the 5' domain confirmed the prediction of a **two stem-loop structure** (Fig. 5 and Fig. 6/A). Note that since neither residue U49 nor residue U60 was modified by CMCT (Fig. 5/A), and that the bond G48-U49 was cleaved by RNase V1 (Fig. 5/B), these two residues may be hydrogen bonded in stem Ib (Fig. 6/A).

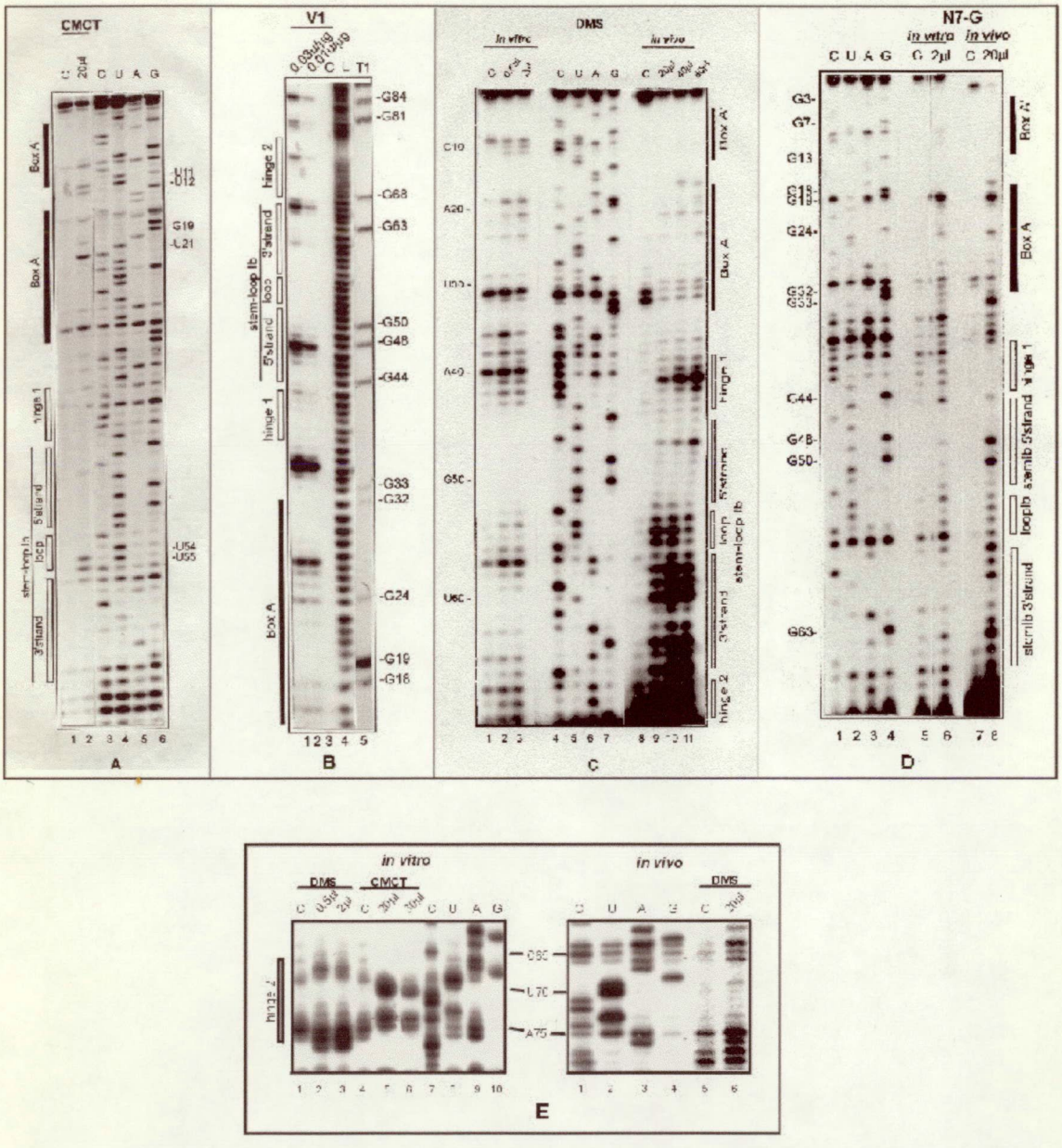


**Figure 5.** Chemical and enzymatic probing of the 5' domain (from nt 1 to 70) of *Cre* U3 snoRNA in solution (Panels A, B, C and D, Lanes marked *in vitro*), and DMS modification of *Cre* U3 snoRNA *in vivo* (Panels C and D, Lanes marked *in vivo*). RNA samples and *Cre* cells were treated by chemical reagents and enzymes as described in Materials and Methods. Positions of cleavages and modifications were identified by extension of primer P/Internal Loop (see Materials and Methods) with reverse transcriptase. The cDNAs were fractionated on a 7% sequencing gel. The amounts of chemical reagents (DMS, CMCT) and enzymes (RNases V1 and T1) used are indicated on the top of the figure. Control Lanes (C) correspond to primer extensions made on RNA incubated in the absence of chemical reagent or enzyme. Lanes C, U, A, G correspond to the sequencing ladder. Nucleotide positions in *Cre* U3 and secondary structural elements are marked on the left and right side of the panels.

Compared to earlier models for higher plant U3 snoRNAs (Kiss and Solymosy, 1990; Marshallsay et al., 1990), *Cre* U3 snoRNA shows **extended hinges 1 and 2** (Fig. 6/A). Similar extensions of these two hinges were also found in yeast U3 snoRNA (Méreau et al., 1997). To test for possible structural changes of U3 snoRNA *in vivo, Cre* cells were exposed to DMS and the methylation pattern of U3 snoRNA was analyzed by primer extension as described in Materials and Methods (Fig. 5/C, D, E). The results are summarized in Fig. 6/B and a **model of interaction with the 18S part** of the pre-rRNA is proposed in Fig. 6/C and D. Residues A9 and C10, which are constituents of box A', and residue A20, a constituent of box A, were protected *in vivo* but not *in vitro*. This is in agreement with the involvement of residues A9, C10 and A20 in the bimolecular helices III and II, respectively (Fig. 6/C). In contrast to their behavior *in vitro,* residues A14, and particularly A15, that are both located between Box A' and Box A, were highly available *in vivo* (Fig. 6/B). This fits with their location between two heterologous helices *in vivo* (Fig. 6/C). Interestingly, in the region from position 1 to position 28 only residues G13 and G24 were found to have an increased sensitivity to DMS at position N-7 *in vivo*. They are both located in a single-stranded segment joining two intermolecular helices. Residue A29 was not reactive to DMS *in vivo,* whereas an increased reactivity was observed for residue A36. This may be explained by the formation of a fourth bimolecular helix IV, as proposed in Fig. 6/C. Note, that residue G34 of this putative helix IV was highly methylated *in vivo* at position N-7.

In hinge 1, residues C41 and C42 were highly protected *in vivo* as compared to *in vitro* (Fig. 5/C and 6/B). In the region corresponding to the 3' strand of stem Ib (Fig. 6/B), all the A and C residues, located between positions A57 and C65, showed a very strong increase in reactivity *in vivo* (Fig. 5/C). This strongly suggests that **stem Ib is not present *in vivo*** and that the portion of the *Cre* U3 snoRNA between positions 53 and 67 is single-stranded. Owing to the limited number of A and C residues in the segment corresponding to the 5' strand of stem Ib, *in vivo* DMS modification gave limited information on this region. The only C residue of this strand (C46) was much more reactive *in vivo* than *in vitro* (Fig. 5/C).

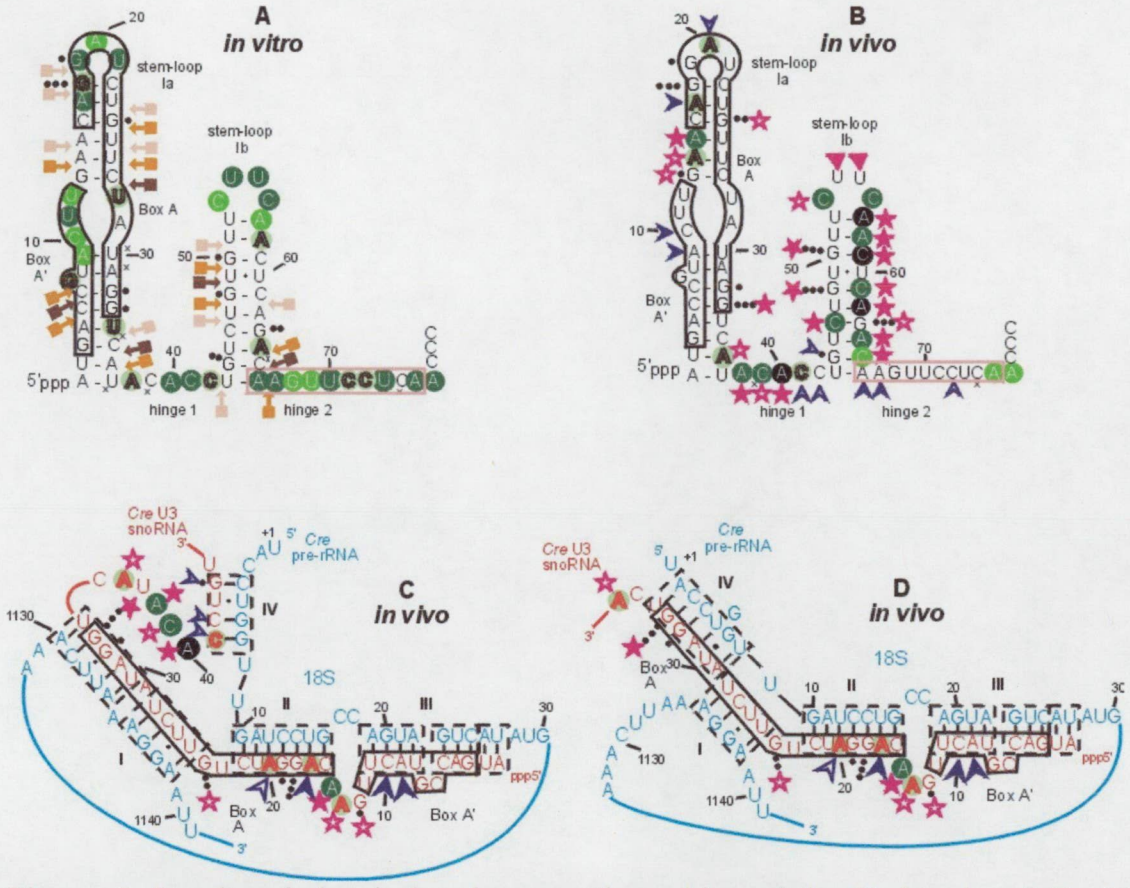


**Figure 6.** Results of chemical and enzymatic probing of the 5' domain of *Cre* U3 snoRNA.

(A) Proposed secondary structure of the 5' domain of *Cre* U3 snoRNA as deduced from structure probing experiments *in vitro*: the phylogenetically conserved boxes A' and A are framed in black. Nucleotides modified by DMS or CMCT at Watson-Crick positions are circled. N7-G methylations are shown by black dots, the number of dots reflects the intensity of the modification. Positions of RNase V1 cleavages are indicated by arrows linked to squares. Intensity of the colors (green for chemicals, orange for V1) indicates the intensity of the modification or yield of cleavage. Crosses (x) indicate pauses of the reverse transcriptase.

(B)Reactivity of nucleotide residues to DMS *in vivo* and changes in sensitivity under *in vivo* as compared to *in vitro* conditions. For easy comparison the RNA is drawn as in Panel (A), although the *in vitro* secondary structure certainly does not exist in the cell. Code for DMS modification as in Panel (A). In addition, extremely strong modifications are shown in black. Decreased sensitivity to DMS indicated by blue arrow: open arrow for a moderate protection, full arrow for a strong protection. Increased sensitivity is indicated by purple star: open star for a moderate increase, full star for a strong increase. Unusual modifications of U residues are indicated by purple triangles (see text for explanation). The hinge 2 sequence proposed to be complementary to the 5'ETS rRNA is framed in pink.

(C) and (D) Two alternative models of the *Cre* U3 snoRNA – pre-rRNA interaction. Bimolecular helices I, II, III and IV between the *Cre* U3 snoRNA (sequence in red) and the *Cre* 18S rRNA (sequence in blue) according to Hughes (15) and Méreau *et al*, (14) are shown. The phylogenetically conserved boxes A and A' are boxed in black. Variation of reactivity of nucleotides to DMS under *in vivo*, as compared to *in vitro* conditions, are indicated with the same code as in Panel (B).

## 4.2 THE 3' DOMAIN OF *CRE* U3 SNORNA

### 4.2.1 Secondary structure of the 3' domain of *Cre* U3 snoRNA in solution

A summary of the results of structure probing reactions of *Cre* U3 snoRNA in solution and a predicted two-dimensional structure of the RNA, called the Wild-Type 1 (WT-1) structure, are shown in Fig. 7.

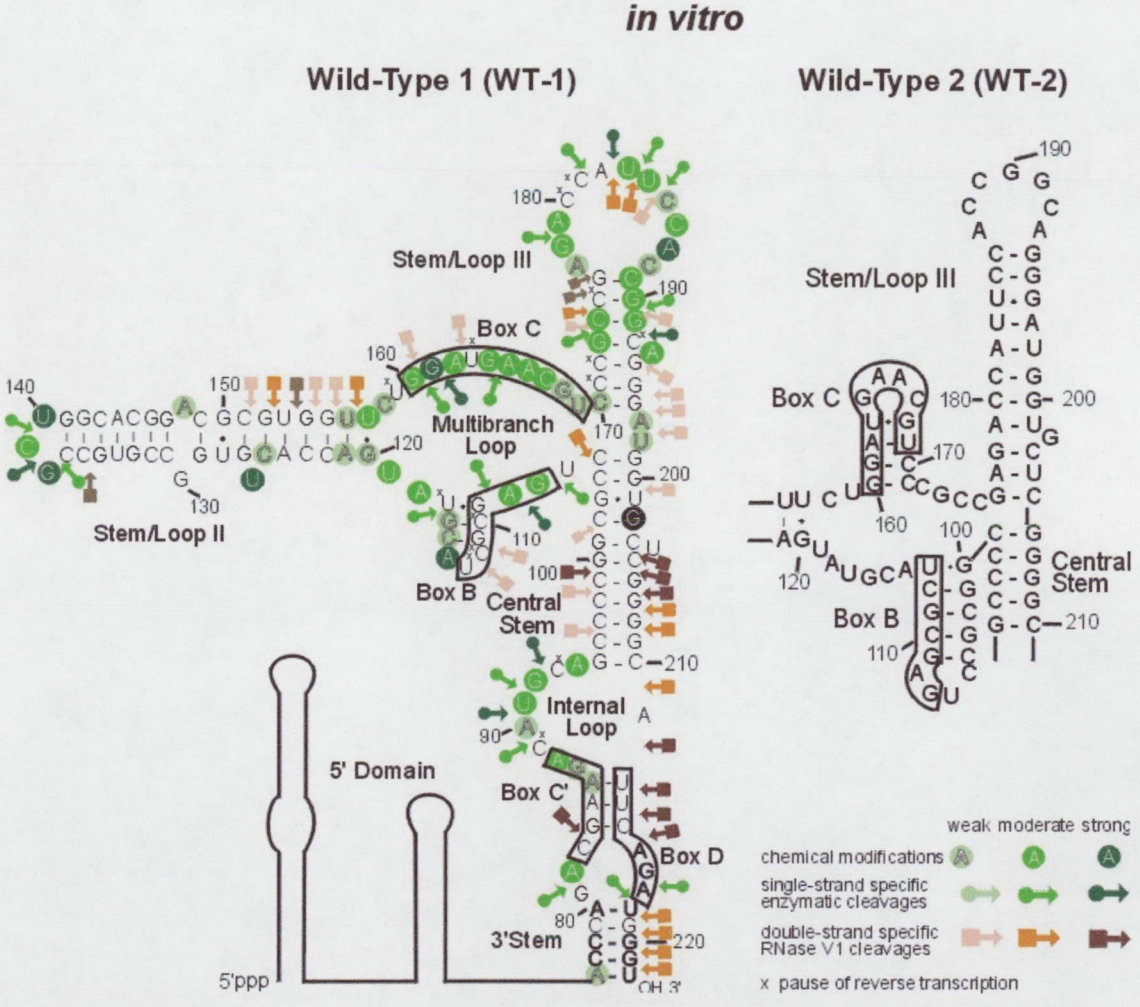


**Figure 7.** Summary of the results of the *in vitro* structure probing experiments of the 3' domain of the wild-type *Cre* U3 snoRNA. Data are drawn on a structure derived from computer prediction (Wild-Type 1). The phylogenetically conserved boxes B to D are framed. Nucleotides modified by chemical probes are circled in green, and sites of single-strand- and double-strand specific enzymatic cleavages are marked by green arrows with circular and by pink, red and brown arrows with quadratic tails, respectively. Color intensity indicates strength of modification or cleavage. X marks nucleotides that could not be evaluated owing to pause of reverse transcription. Symbols in green refer to single-stranded, and those in reddish to double-stranded regions. Wild-Type 2 structure shows the alternative folding possibility of the 3'domain of *Cre* U3 snoRNA.

The 3' domain of *Cre* U3 snoRNA is particularly rich in G-C base-pairs that constitute four helices, called Stem II, Stem III, Central Stem and 3' Stem, that are bracketing the Multibranch Loop and the Internal Loop regions (Fig. 7). The long GC-rich helices all over the RNA result in regions with strong secondary structure, in which neighboring regions were hard to analyze.

In the **Multibranch Loop** region which accommodates the conserved **boxes B** and **C**, nucleotides G109-C112 of box B were fully resistant to single strand-specific chemical reagents (Fig. 8/B, lanes 1, 2, 3) and RNases, whereas RNase V1 cleavages were detected at G111 and C112 (Fig. 7).

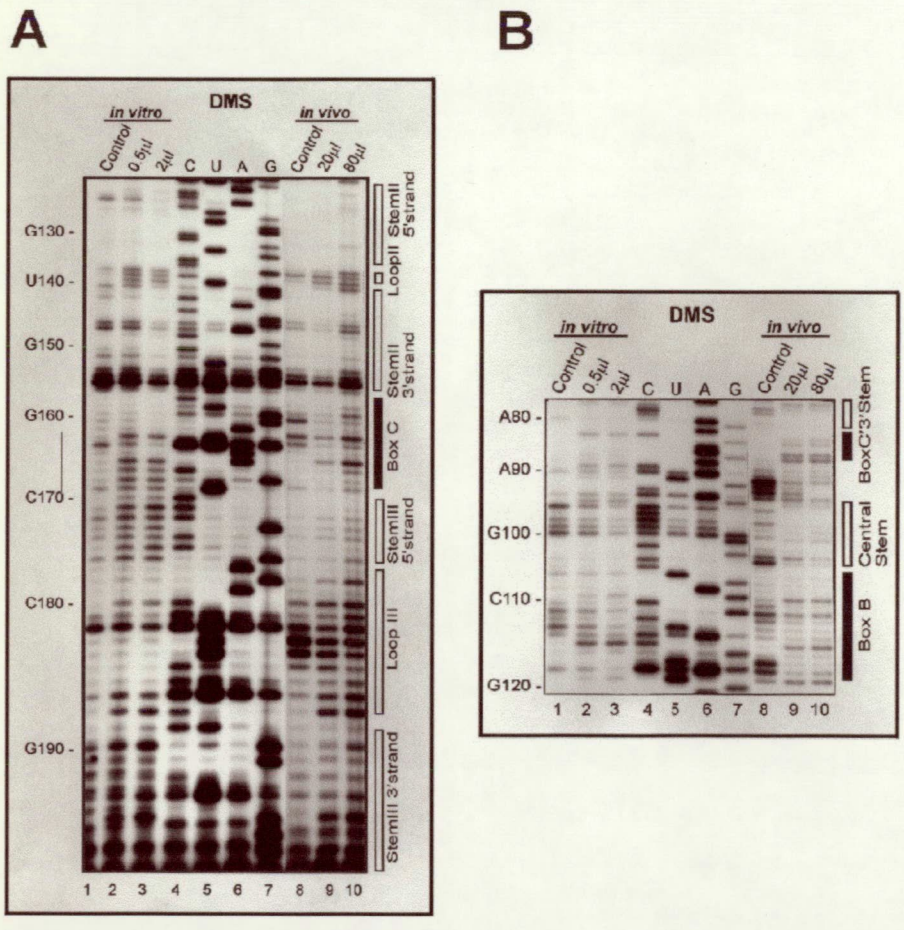


**Figure 8.** Chemical modification by DMS of the 3' domain (from nt 190 to nt 120 [panel A] and from nt 120 to nt 80 [panel B]) of *Cre* U3 snoRNA *in vivo* (lanes 8-10). RNA samples were treated as described in Materials and Methods, reverse transcribed using primers P/3' (see Materials and Methods) and loaded onto 7% sequencing gel. The amounts of DMS are indicated at the top of the figure. Control: incubation of the RNA in the absence of DMS. Dideoxynucleotide sequencing lanes are denoted according to the base of the RNA they detect. Nucleotide positions and secondary structural elements are marked on the left and on the right, respectively. To assess changes in DMS modification pattern of the molecule under *in vivo* relative to *in vitro* conditions, accessibility to DMS of *Cre* U3 snoRNA in solution was investigated in parallel (lanes 1-3).

These observations suggest that the G109-G111 region likely forms a base-pairing interaction with the C115-U117 region. In contrast to what has been found with the yeast U3 snoRNA (Méreau et al., 1997), neither single strand-specific probes nor RNase V1 provided evidence for an interaction between residues A166-U169 of box C and A114-U117 of box B. In the Internal Loop, RNase V1 cleavages showed that nucleotides G84-A86 in box C' are base-paired with nucleotides U212-C214 in box D.

Our results of structure probing of *Cre* U3 snoRNA fully confirmed the existence of Stem/Loop II, the Central Stem and the 3' Stem. However, the Stem/Loop III region exhibited equivocal reaction patterns towards single and double strand-specific agents (Fig. 9, lanes 5-14). In Stem III, although several nucleotides were cleaved by RNase V1 (G173-G176, G190, A193, G195-A196) (Fig. 9, lanes 6 and 7), residues G173, G190 and G191 were extensively modified by kethoxal (Fig. 9, lanes 12-14) and hydrolyzed by RNase T1 (not shown). In the terminal Loop III region, residues A182, U183 and U184 frequently exhibited strong RNase V1 cleavages (Fig. 9, lane 7). Moreover, nucleotides G168-C172 and C174-C175 were sensitive to chemical modifications. These observations indicate that, beside the WT-1 structure the Stem/Loop III region of *Cre* U3 snoRNA can adopt an alternative two-dimensional structure in solution.

Computer-aided RNA folding analyses revealed that the Stem/Loop III and the Multibranch domains of the *Cre* U3 snoRNA can be folded into another structure, called the Wild-Type 2 (WT-2) structure (Fig. 7). The WT-2 structure is fully consistent with all those structure-probing modifications obtained by us, which do not conform to the WT-1 structure. In the WT-2 RNA, Stem/Loop III, the Multibranch Loop and part of the Central Stem are completely rearranged as compared to the WT-1 RNA. The Central Stem of the WT-2 RNA is 6 base pairs shorter than that of the WT-1 isomer, since nucleotides G100-C104 form of 5 base-pairs helix with nucleotides G109-U103 in the box B region. Nucleotides G199-C205 that are located in the 3' strand of the Central Stem of the WT-1 RNA become part of Stem III in the WT-2 structure. The bulged G201 residue is an excellent marker for the WT-2 *Cre* U3 snoRNA, since it is readily detectable by kethoxal (Fig. 9, lanes 13 and 14). In the WT-2 RNA, residues G190 and G191 reside in Loop III and G173 is part of a single-stranded region that connects box C and Stem III. Nucleotides from A182 to C185 become part of the 5' strand of Stem III. Based on the intensity of distinctive structure probing reactions, the two structural

isomers of the *Cre* U3 snoRNA coexist in approximately equimolar amounts in solution. Alteration of $Mg^{2+}$ concentration or denaturation and renaturation failed to stabilize any of the two structures (data not shown).

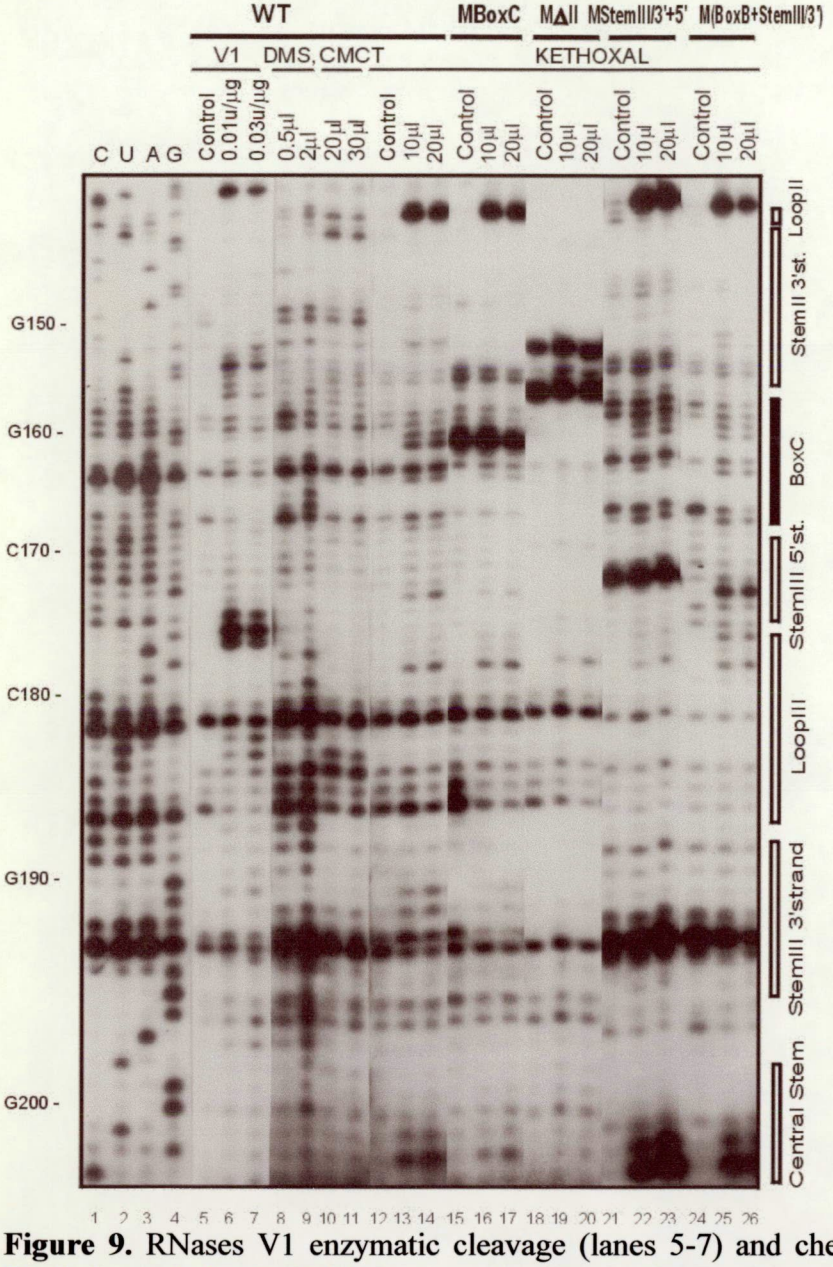


**Figure 9.** RNases V1 enzymatic cleavage (lanes 5-7) and chemical modifications by DMS (lanes 8, 9), CMCT (10,11) and kethoxal (12-14) of a part (from nt 203 to nt 135) of the 3' domain of the Wild-Type (WT) *Cre* U3 snoRNA in solution. For the comparison of the changes in the structure the samples of kethoxal modifications of different mutants are loaded parallel (lanes 15-26). RNA samples were treated as described in Materials and Methods, reverse transcribed using primer P/3'Eco (see Materials and Methods) and separated on a 7% sequencing gel. The amounts of the chemicals and enzyme are denoted at the top of the figure. Control: incubation of the RNA without reagent. Dideoxynucleotide sequencing lanes are denoted according to the base of the RNA they detect. Nucleotide positions and secondary structural elements are marked on the left and on the right, respectively. Reverse transcriptase stops one nucleotide prior to a modified base.

### 4.2.2 Mutations stabilizing the Wild-Type 1 and 2 structural isomers of *Cre* U3 snoRNA

To define structural elements that are responsible for the instability of Stem/Loop III, a series of substitution (M) and deletion (MΔ) mutants of the *Cre* U3 snoRNA were constructed and subjected to *in vitro* structure probing. As a first step, all mutant RNAs were probed with lead-acetate and those which exhibited a mixed cleavage pattern comparable to the wild-type *Cre* U3 snoRNA were excluded from further analyses. The experimental data of the structure probing of four mutant *Cre* U3 RNAs that preferentially adopted the WT-1 (MBoxC and MΔII) or the WT-2 [StemIII/3'+5' and M(BoxB+StemIII/3')] structure are summarized in Fig. 10. The results of structure probing experiments performed with kethoxal are shown in Fig. 9 (lanes 12-26). The WT-1 structural conformation of the *Cre* U3 snoRNA features highly reactive G residues at positions 173 and 202, while the reactive G or U residues at positions 190 and 191 are characteristic of the WT-2 structure. It is noteworthy that none of the tested mutations affected the two-dimensional structure of the 5' domain (from A1 to U69) of *Cre* U3 RNA (data not shown).

In the **MBoxC** mutant (Fig. 10/A), replacement of G160 and G161 in the C box motif with U residues resulted in a decreased accessibility to kethoxal of residues G173, G190, G191 and G202 (Fig. 9, lanes 15-17). Parallel nucleotides located in the Loop III region showed an increased reactivity with single strand-specific chemical and enzymatic probes and, contrary to the wild-type *Cre* U3 snoRNA, no RNase V1 cleavages were observed in this region. These effects were even more conspicuous when the entire Multibranch Loop together with the Stem/Loop II structure was deleted from the *Cre* U3 snoRNA. In the resulting **MΔII** RNA (Fig. 10/B), residues from G178 to C185 showed an increased accessibility to single strand-specific RNA modifying agents (Fig. 11/B), and no RNase V1 modification was detectable in Loop III (Fig. 11/A). Residues predicted to form base pairs in the Stem III and Central Stem regions of the MΔII RNA reacted exclusively with the double strand-specific RNase V1.

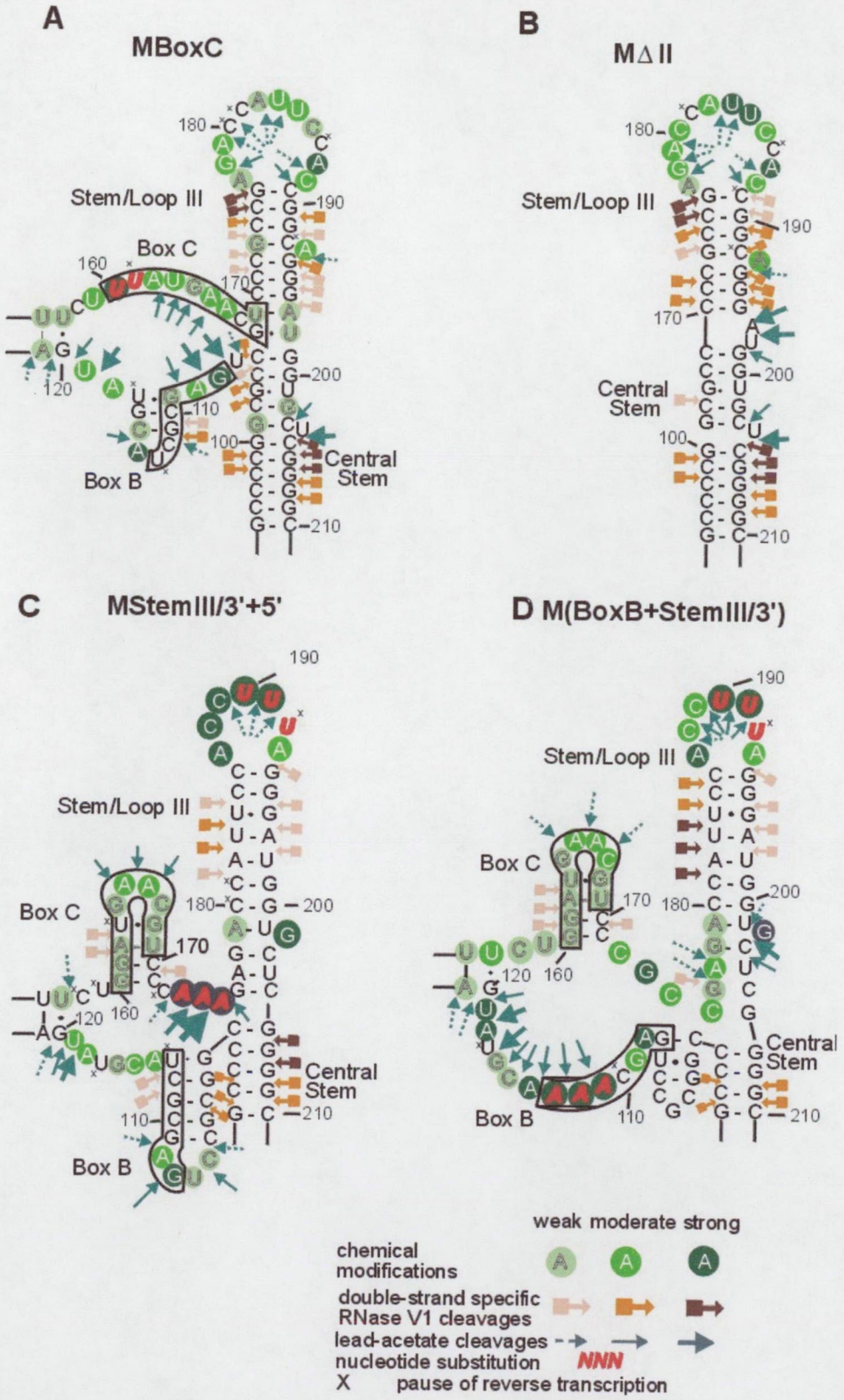


**Figure 10.** Alternative folding patterns of the 3'domain of *Cre* U3 snoRNA mutants as suggested by to the Wild-Type 1 or the Wild-Type 2 model (E). The mutated nucleotides are in red. The different mutants - MBoxC (C), MΔII (D), MStemIII/3'+5' (F), and M(BoxB+StemIII/3') (G) - are drawn in a structure suggested by the data of the structure probing experiments. The symbols are identical with those explained in Fig. 6.

Taken together, the above results demonstrate that the Multibranch Loop and the Stem/Loop II regions play a fundamental role in the structural destabilization of Stem/Loop III and of the Central Stem of *Cre* U3 snoRNA. Of course, an interaction between the Multibranch Loop and the neighboring stem structures is expected to be reciprocal. Therefore, it was not surprising to find that upon removal of Stem/Loop III (from C170 to U198), the Multibranch Loop lost its single-stranded nature and novel base-pairing interactions were formed between nucleotides A166-U169 and U117-A114 as well as nucleotides U113-A108 and G199-U204 (data not shown).

As predictable, after destabilization of the Stem III helix of WT-1 RNA, either by replacement of the 190-GGC-192 trinucleotides with three U residues or the 173-GCC-175 fragment with three A residues, the WT-2 structure became preferentially detectable in solution (data not shown). More surprisingly, a combination of the above mutations in the **MStemIII/3'+5'** RNA, although it was expected to restore the Stem III structure of the WT-1 RNA, still showed a modification pattern characteristic of WT-2 RNA (Fig. 10/C). The newly introduced U190, U191, A173, A174 and A175 residues were highly accessible to lead-acetate and kethoxal, demonstrating that replacement of three G-C base-pairs with three A-U base-pairs in Stem III of the WT-1 RNA could already destabilize this helix and facilitate preferential formation of the WT-2 structure.

It has been shown above that a short hairpin formed within the box C region plays an essential role in the stabilization of the WT-2 structure. We therefore tested the importance of another short hairpin of the WT-2 RNA formed by the G109-U113 residues in the box B and the G100-C104 sequences. Since the latter nucleotides constitute the upper part of the Central Stem of the WT-1 RNA, one can assume that the hairpin in the box B region may contribute to the observed unstable behavior of the Central Stem of *Cre*, yeast and human U3 snoRNAs (Méreau et al., 1997; Parker and Steitz, 1987). To prevent formation of this hairpin, the residues G111, C112 and U113 of *Cre* U3 RNA were substituted by residues AAA. Structure probing of the resulting **MBoxB** RNA, however, showed a mixed wild-type modification pattern (data not shown), demonstrating that this hairpin, contrary to the box C-type hairpin, cannot alone induce conformational changes in *Cre* U3 snoRNA. To assay how the Central Stem is folded in the WT-2 RNA when formation of the „box B hairpin" is inhibited, the GCU to AAA mutations present of the MBoxB RNA were introduced into the StemIII/3' RNA. The resulting **M(BoxB+StemIII/3')** RNA showed an extremely high accessibility

at residue G202 (Fig. 9, lanes 24-26), indicating that nucleotides from U201-G206 did not form the eleven base-pair Central Stem that would characteristic of WT-1 RNA. The strong RNase V1 cleavages observed at the residues C180, C181, A182, U183, U184, C185 and C186, together with a complete protection of residues U183 and U184 against CMCT, provided strong arguments that the M(BoxB+StemIII/3') RNA adopted exclusively the WT-2 structure.

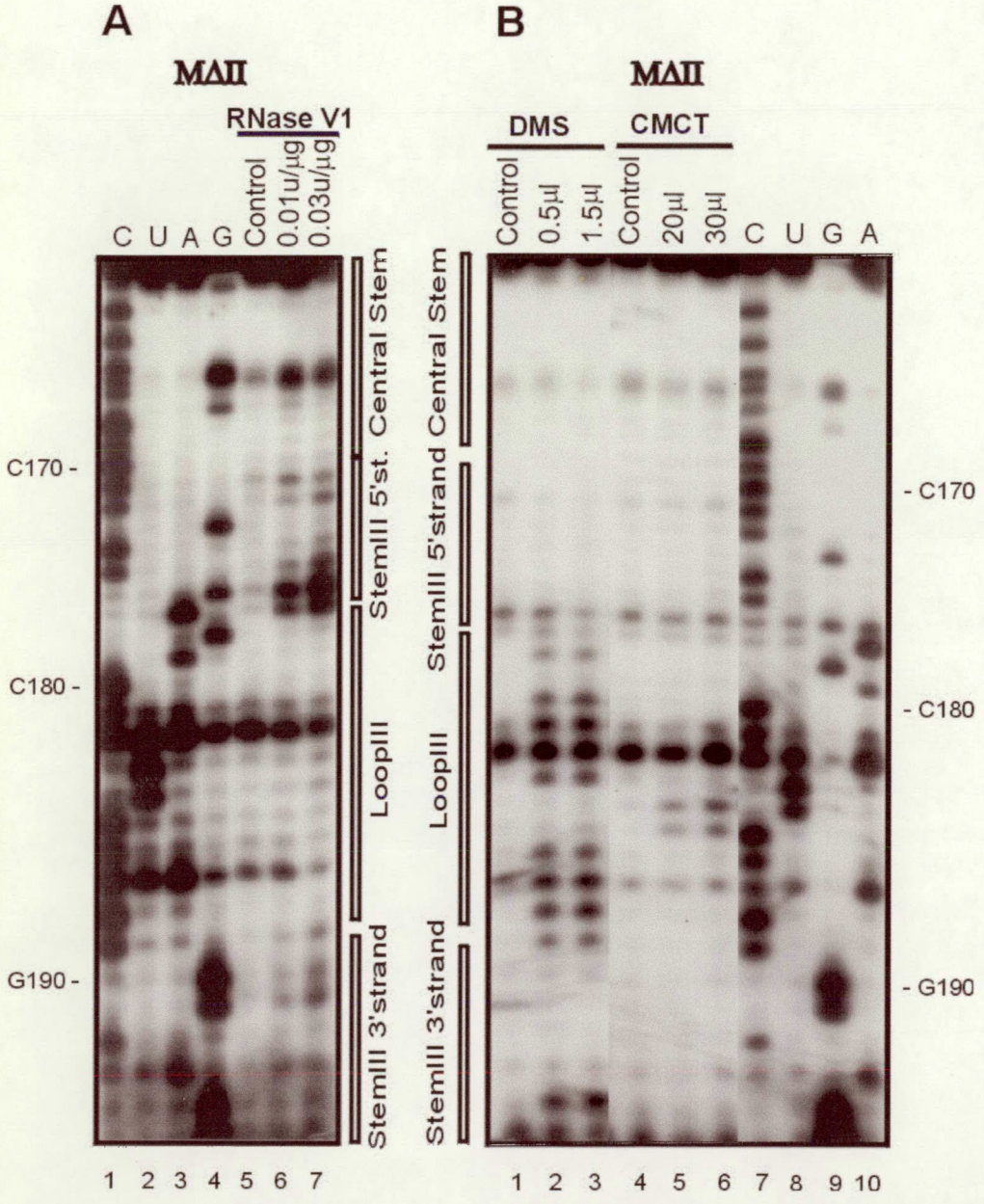


**Figure 11.** Structure probing analyses of the MΔII mutant of *Cre* U3 snoRNA by RNase V1 (A, lanes 5-7), DMS (B, lanes 1-3) and CMCT (B, lanes 4-6) using primer P/3' (see Materials and Methods) for extension. (For the details see the legend to Fig. 9).

### 4.2.3 The 3' domain of *Cre* U3 snoRNA may form a Wild-Type 1 structure *in vivo*

To get an insights into the two-dimensional structure of the *Cre* U3 snoRNA under *in vivo* conditions, intact *Cre* cells were exposed to DMS, total RNA was isolated and the reactive nucleotides were monitored by primer extension using terminally labelled U3-specific oligodeoxynucleotides (Fig. 8). The results of the *in vivo* structure probing of *Cre* U3 are summarized in Fig. 6. Unfortunately, the G residues, the most helpful markers of the two isomers of *Cre* U3 snoRNA, cannot be analyzed by DMS. Therefore, all conclusions for the *in vivo* conformation of the *Cre* U3 snoRNA were based on the A and C modification patterns.

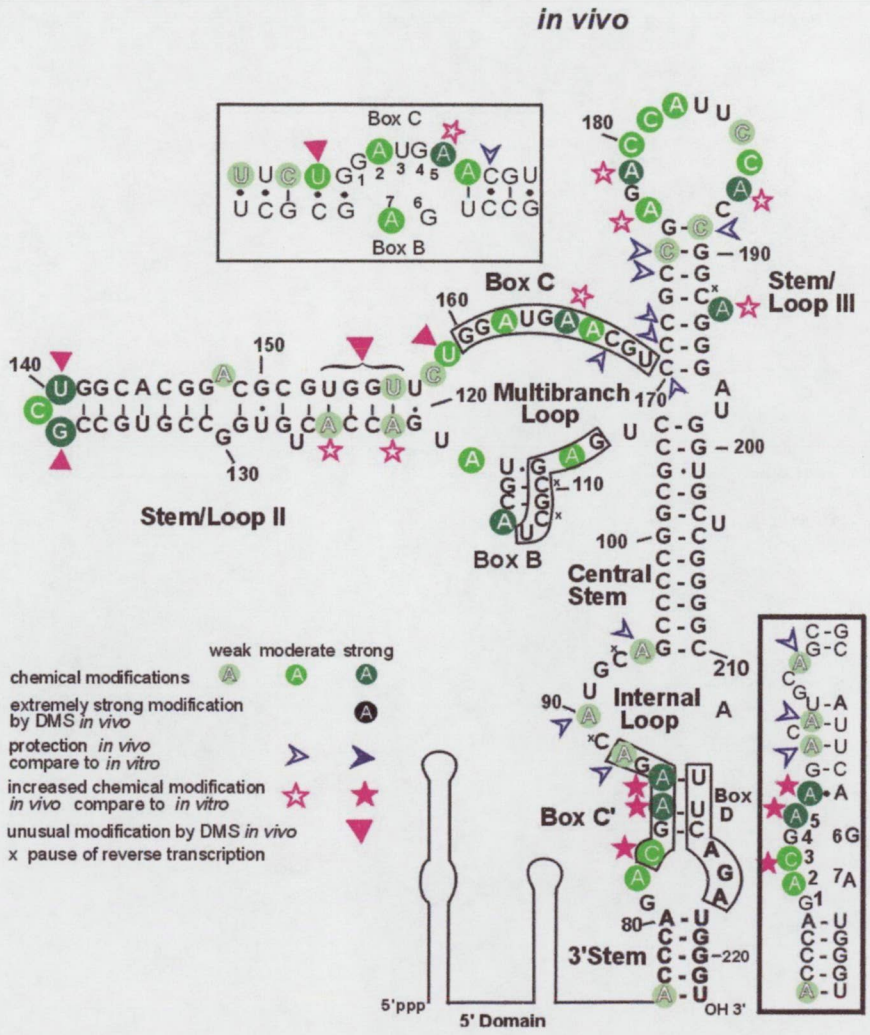


**Figure 12.** Summary of the results of *in vivo* modification by DMS of the 3' domain of *Cre* U3 snoRNA (RNP). The results are shown on the secondary structure drawn according to the Wild-Type 1 model of U3 snoRNA as established under *in vitro* conditions. Nucleotides modified by DMS are circled in green. Changes in sensitivity to DMS under *in vivo* relative to *in vitro* conditions are represented by arrowheads and stars: blue arrowhead, protection (filled, strong; open, moderate); violet star, increase in sensitivity (filled, strong; open, moderate).

In the 3' domain of *Cre* U3 snoRNA, only the 3' Stem and the box B region behaved identically under *in vitro* and *in vivo* conditions. In the box B region, the three A residues were equally accessible to DMS, but none of the C residues was modified (Fig. 8/B). This is consistent with the presence of a short helix formed between residues G109-G111 and C115-U117, formed in WT-1 RNA. The Stem II region, although it was highly structured *in vitro,* appeared to be more relaxed *in vivo* (Fig. 8/A), since both strands of the Multibranch Loop-proximal helix underwent U and A modifications. A similar phenomenon was observed for the yeast U3 snoRNA (Méreau et al., 1997), in which some A residues in Stem/Loop II became highly accessible to DMS under *in vivo* conditions. Consistent with this observation, both in the human and yeast U3 snoRNAs (Méreau et al., 1997; Parker and Steitz, 1987), helices located next to the Multibranch Loop region were found to be resistant to RNase V1 digestion in the U3 snoRNP. Therefore, we suggest that snoRNP proteins bound directly to this helix structure and/or to the neighboring Multibranch Loop region might destabilize this helix of U3 snoRNA.

Since no chemical modification was detected in the Central Stem (Fig. 8/B), the structure of this region can only be indirectly inferred from the structure probing of other parts of the *Cre* U3 snoRNA. In the Stem III region, C175 represents the only accessible residue and contrary to the *in vitro* structures, the C residues at positions 170-172 and 174 are fully protected in *Cre* cells (Fig. 8/A). The unpaired nature of nucleotides A179, C180, C181, C185 and C186 located in the Loop III region was more explicit in the U3 snoRNP than in the naked U3 RNA. The box C region showed very similar *in vitro* and *in vivo* DMS modification patterns, although under *in vivo* conditions residue C167 became resistant to DMS and residue A165 showed an increased sensitivity in the cell (Fig. 8/A). Under some conditions, DMS treatment is known to produce unusual uracil modifications (Ajuh and Maden, 1994). The exceptional *in vivo* uracil modifications at U184 and U159 confirmed the single-stranded nature of both the Loop III and box C regions in *Cre* U3 snoRNA.

Residue C167 in box C of *Cre* U3 snoRNA showed no DMS modification in *Cre* cells, contrary to the *in vitro* RNA modification reactions. The equivalent C residue in the yeast U3 snoRNA, C259 that is conserved in all known U3 snoRNAs (Solymosy and Pollák, 1993), was also protected *in vivo* against single strand-specific chemical probes and the double strand-specific RNase V1 (Méreau et al., 1997). Structure probing of human (Parker and Steitz, 1987) and *X. laevis* (Jeppesen et al., 1988) U3

snoRNP showed that the conserved residues C165 and C164, respectively, were resistant to chemical probes. These data strongly support the view that the conserved residue C167 in the box C region of the *Cre* U3 snoRNA might directly interact with snoRNP protein(s).

In the box C' region, the increased sensitivity of residues C83, A85 and A86 to DMS may indicate that the base-pairing interaction between boxes C' and D observed in solution is disrupted in the U3 snoRNP by protein factors that bind to this regions (Fig. 8/B). The decreased reactivity of residue A90 suggests that this nucleotide may also contribute to the snoRNA and proteins interaction. The opposite strand of the Internal Loop region, including the box D motif, could not be analyzed, since these sequences served as a docking site for the 3'-terminal primer.

In summary, we may conclude that under *in vivo* conditions the 3' domain of *Cre* U3 snoRNA folds into a two-dimensional structure that is highly reminiscent of the WT-1 structure. Most likely, snoRNP proteins bound to the U3 RNA inhibit formation of the alternative WT-2 structure and stabilize the WT-1 (or a WT-1-like) structural conformation in the nucleoli of *Cre* cells.

# 5 DISCUSSION

A detailed comparison of the U3 snoRNA modification patterns obtained in solution and in *Cre* cells allowed us to identify the two-dimensional structure of and the potential protein binding sites in *Cre* U3 snoRNA.

## 5.1 The 5' terminal domain of *Cre* U3 snoRNA forms several intermolecular helices with the 18S pre-rRNA *in vivo*

The 5' domain of U3 snoRNA was found to play an essential role in the function of U3 snoRNA (Sharma and Tollervey, 1999). Base-pair interactions with pre-rRNA seem to be required for cleavages of pre-rRNA at sites A0, A1 and A2 in yeast (Beltrame and Tollervey, 1995)(Fig. 1), and at sites 1, 2, 3 in *X. laevis* (Borovjagin and Gerbi, 1999). Interestingly, the 5' domain of vertebrate and trypanosome U3 snoRNA folds into a single stem-loop structure (Parker and Steitz, 1987; Jeppesen et al., 1988; Hartshorne and Agabian, 1994), whereas that of plant U3 snoRNA and of U3 snoRNAs from several lower eukaryotes (yeasts, ciliate) can be folded into a two stem-loop structure (Porter et al., 1988; Kiss and Solymosy, 1990; Marshallsay et al., 1990), for review: (Solymosy and Pollák, 1993). Experimental demonstration of this two stem-loop structure was obtained only for yeasts (Ségault et al., 1992). Here, we show by chemical and enzymatic probing of the 5'-terminal domain of an *in vitro* synthesized *Cre* U3 snoRNA that in *Cre* too, it is folded in a two stem-loop structure (Fig. 6/A). Stem-loop structure Ia contains the two phylogenetically highly conserved boxes A and A' (Wise and Weiner, 1980; Myslinski et al., 1990). Interestingly, a U.U pair is part of helix Ib, since no modification by CMCT of the two U residues is observed, and the bordering phosphodiester bonds are cleaved by V1 RNase. As previously found in *S. cerevisiae* (Méreau et al., 1997), the DMS modification pattern of the 5' domain of *Cre* U3 snoRNA was completely different under *in vitro* (Fig. 6/A) and *in vivo* (Fig. 6/B) conditions. The data obtained *in vivo* strongly support the formation of bimolecular helices I and II (Fig. 6/C and 6/D) between box A of U3 snoRNA and the 5'-terminal region of the 18S rRNA part of pre-rRNA. The *Cre* U3 nucleotides expected to be involved in the formation of **helix I** or **II** showed DMS protection *in vivo*, in contrast to their behavior under *in vitro* conditions, whereas nucleotides in linker segments showed an increased reactivity. Utilization of DMS *in vivo* to probe the *S. cerevisiae* U3

snoRNA structure led Méreau *et al.* (Méreau et al., 1997) to suggest modifications of the U3 - pre-rRNA interaction model (Fig. 2/A) proposed by Hughes (Hughes, 1996). This included a reduction of the size of helix II. In *Cre* cells we confirm the 7bp length of helix II. Indeed, the increased reactivity of residue $G_{24}$ in *Cre* U3 *in vivo* shows that this residue is not base-paired with the 18S part of pre-rRNA. Sharma and Tollervey (Sharma and Tollervey, 1999) recently demonstrated the importance of formation of helices I and II for cleavages at sites A1 and A2 in *S. cerevisiae*. According to previous data on the *S. cerevisiae* U3 snoRNA an extension of helix III was proposed (Méreau et al., 1997). Based on sequence comparisons, such an extension is possible in several lower eukaryotes, and in plants, but not in vertebrates (Méreau et al., 1997). The nucleotide sequence of *Cre* U3 snoRNA shows that a 3 bp extension of **helix III** similar to that found in yeast can be formed in *C. reinhardtii*. In view of the absence of DMS reactivity of the 5 residues at the 5' terminus of *Cre* U3 snoRNA *in vivo*, formation of an extended helix III in *C. reinhardtii* is very likely, although we have no direct proof of its formation, since the five 5'-terminal residues of *Cre* U3 snoRNA are also resistant to DMS *in vitro*. However, the sequence from position 33 to position 37, which is the partner of the terminal sequence in the *Cre* U3 snoRNA structure found *in vitro*, has a modified reactivity *in vivo* (Fig. 6B): high reactivities to DMS of the N-1 position of residue $A_{36}$ and the N-7 position of residue $G_{33}$ are detected *in vivo*. This is a strong indication for a conformational change of this RNA region *in vivo*. In other words, its base pairing with the 5' extremity of *Cre* U3 snoRNA *in vivo* is very unlikely. This is in favor of an extended helix III. Finally, formation of a fourth helix (IV) between U3 snoRNA and the 18S part of pre-rRNA, was proposed in *S. cerevisiae* (Méreau et al., 1997). The *S. cerevisiae* U3 snoRNA sequence involved in this base-pair interaction is located downstream of the segment that base pairs with the 5'-ETS region (Fig. 2/A). Such complementarity with the 18S part of pre-rRNA is not well conserved in *C. reinhardtii*. Since residues $A_{29}$ and $A_{31}$ were not modified under *in vivo* conditions, we propose the formation of a **fourth helix (IV)** between a stretch (nt 29 to nt 34) of *Cre* U3 snoRNA and the 18S part of pre-rRNA (Fig. 6/D). Additional experiments are needed to confirm the existence of this base-pair interaction, since none of the A residues implicated in this segment is modified *in vitro*, and since the N-7 position of

residue G33 included in this putative heterologous helix showed an increased reactivity to DMS.

As already found in yeast, the single-stranded *Cre* U3 snoRNA residues located between the bimolecular helices are accessible to DMS *in vivo*. This suggests the absence of protection by proteins. In *S. cerevisiae*, the Mpp10 protein was proposed to interact with the 5'-terminal domain of U3 snoRNA (Lee and Baserga, 1997), and to play an important role in U3 snoRNA function (Baserga et al., 1997; Lee and Baserga, 1997). As evidenced by the DMS reactivity of U3 snoRNA in *S. cerevisiae* and in *C. reinhardtii in vivo*, either Mpp10 and its possible counterpart in *C. reihardtii* exhibit a transitory interaction with the 5'-terminal domain of U3 snoRNA, that does not result in the protection of the single-stranded regions involved, or these proteins interact with bimolecular helices.

**5.2 The *Cre* U3 snoRNA does not form stem-loop 1b *in vivo*, instead it may interact with the 5'-ETS region of pre-rRNA**

Our *in vivo* probing of *Cre* U3 snoRNA with DMS (Fig. 6/B) revealed strong conformational changes of the portion of *Cre* U3 snoRNA that form stem-loop structure Ib *in vitro* (Fig. 6/A). Clearly, this stem-loop structure is not formed *in vivo*. Based on its very high sensitivity to DMS the segment from position 57 to 65 is single-stranded and devoid of associated protein *in vivo*. A protection of part of hinge 1 and of the N-7 position of residue G44 is detected *in vivo*. Hinge 1 of *S. cerevisiae* U3 snoRNA was found to be involved in base-pair interaction with the 5'-ETS region (Beltrame and Tollervey, 1992). This interaction was necessary for cleavages at sites A0, A1 and A2 of pre-rRNA (Beltrame and Tollervey, 1995). Similarly, the hinge region of *T. brucei* U3 snoRNA was shown to interact with a 5'-ETS region, adjacent to the processed primary site, giving an additional argument for the role of U3 in its processing (Hartshorne, 1998). In *Cre*, a very likely explanation for the observed protection of the segment from nt 41 to nt 45 (or 41 to 52) (Fig. 6/B) and for the absence of formation of stem-loop 1b in *Cre* U3 snoRNA *in vivo* consists in assuming base-pairing interaction with the 5'-ETS region of pre-rRNA. Unfortunately, no 5'-ETS sequence from *C. reinhardtii* is available to confirm this hypothesis, but, as several 5'-ETS sequences from plants are available, we checked whether some complementarities could be found between the U3 snoRNA

of some plants and the corresponding 5'-ETS sequences. For the three plants investigated [rice (*Oriza sativa* acc.number: X54194), maize (*Zea mays*, acc. number: X03989), and Arabidopsis (*Arabidopsis thaliana*, acc.number: X15550)] at least one common 5'-ETS sequence was found complementary to the five nt U3 sequence equivalent to the *Cre* U3 nt 41 to 45 (Fig. 13). Additional experiments, and especially, the determination of the *Cre* pre-rRNA 5'-ETS region are needed to confirm the existence of such an interaction.


**Figure 13.** Alignment of the 5' portion of U3 snoRNAs from *T. brucei, S. cerevisiae, Z. mays, O. sativa, A. thaliana,* and *C. reinhardtii* showing sequence complementarities with the 5'-ETS regions of pre-rRNAs. Data bank accession numbers of the U3 snoRNA sequences are indicated in the legend to Fig. 1. The nucleotides of phylogenetically conserved boxes GAC, A' and A are shown in green. The potential sites of interaction of U3 snoRNAs with the 5'-ETS regions are highlighted in pink. The complementary 5'-ETS sequences are in italics, and positions of their extremities as referred to the pre-rRNA transcription start site are indicated. The two 5'-ETS sequences, complementary with U3, given for *T. brucei* were proposed by Hartshorne and Tokoyufu (Hartshorne and Toyofuku, 1999), $U_{862}$ from the 5'-ETS shown to be cross-linked to U3 snoRNA is marked by an asterisk. The *S. cerevisiae* 5'-ETS complementarity corresponds to the well-documented helix V (Beltrame and Tollervey, 1992). Positions of helices I, II, III and IV, formed with the 18S part of the pre-rRNA are shown. The U3 – 5'-ETS complementarities proposed for *Z. mays, O. sativa* and *A. thaliana* were obtained by a search performed with the pattern(n) software (Infobiogen). For *C. reinhardtii* (present paper) and for *S. cerevisiae* (Méreau et al., 1997) the changes of DMS reactivities [the increased (purple stars), or decreased (blue arrows) sensitivity of A and C residues to DMS *in vivo*] are represented with the same code as in Fig. 6/B.

Interestingly, an unusual reactivity to DMS of the two U residues 54 and 55 was detected *in vivo*. This may be explained by some protein interactions altering the

electronegativity of the base ring, as recently observed for N-7 of an adenosine residue in Sm binding sites of UsnRNAs (Hartmuth et al., 1999). It should be noted that the observed protections of *Cre* U3 snoRNA *in vivo* are partial as it was also found for the *S. cerevisiae* U3 snoRNA. This suggests that under the growth conditions used, not all *Cre* U3 snoRNA molecules were base-paired with pre-rRNA similar to *S. cerevisiae*. However, the data obtained are clear enough to detect *Cre* U3 snoRNA conformational changes and define more precisely some base-pare interactions during processing events of pre-rRNA.

### 5.3 The 3' domain of *Cre* U3 snoRNA shows two alternative structures in solution but only one *in vivo*: a stabilizing role of proteins

The experimental data derived from the structural analyses of RNA synthesized *in vitro* demonstrated that the 3' domain of *Cre* U3 snoRNA folds into two alternative structures (Fig. 7). These two isomers of the *Cre* U3, named Wild-Type 1 (WT-1) and Wild-Type 2 (WT-2), coexist in about equimolar amounts in solution. Appropriately designed sequence alterations introduced into the Wild-Type *Cre* U3 snoRNA can stabilize either the WT-1 or the WT-2 conformation. While alteration of the box C region (Fig. 10/A) or deletion of the Stem/Loop II and the Multibranch Loop regions (Fig. 10/B), results in the accumulation of the WT-1 structure, mutations introduced into Stem III )Fig. 10/C) and/or the box B region (Fig. 10/D) stabilize the WT-2 structure. The *in vivo* DMS methylation pattern of the *Cre* U3 snoRNA strongly supports the view that the 3' domain of U3 snoRNA exists in a WT-1-like structure in the nucleoli of *Cre* cells (Fig.12). Unfortunately, structure probing with the double strand-specific RNase V1 is not feasible in *Cre* cells. Hence, we cannot confirm that the WT-1 and the *in vivo* structure of *Cre* U3 snoRNA are really identical in all details.

The 3' Stem represented the only region in the 3' domain of *Cre* U3 that behaved identically during the *in vitro* and *in vivo* structure probing reactions (Fig. 7 and 12). This short stem structure is present in all U3 snoRNAs characterized so far. It is essential for stable accumulation of U3 snoRNA (Samarsky and Fournier, 1998; Terns et al., 1995), and facilitates binding of snoRNP proteins to the neighboring box D and C' regions. The phylogenetically conserved boxes C' and D are always found on the opposite strands of an Internal Loop structure that is bracketed by the 3' Stem and the

Central Stem. Boxes C' and D which most likely function as a single protein binding motif (Xia et al., 1997; Watkins et al., 1998b) are essential for the processing, stability and nucleolar targeting of the U3 snoRNA (Terns et al., 1995). RNase V1 cleavages showed that nucleotides in boxes C' and D can form a short helix in *Cre* U3 snoRNA in solution (Fig. 7). In *Cre* cells, the *in vitro* base-paired part of box C' showed an increased accessibility, indicating that the complementary sequences of box D are available for RNA-protein interactions in the U3 snoRNP (Fig. 12). While a very similar phenomenon was observed for the yeast U3 snoRNA and snoRNP (Méreau et al., 1997), base-pairings between boxes C' and D of the human U3 snoRNA seem to be maintained both *in vitro* and *in vivo* (Parker and Steitz, 1987). In the human U3 snoRNP, the box D region was found to be less sensitive of oligodeoxynucleotide-directed RNase H cleavages than the C' region, suggesting that box D and the surrounding sequences are covered by snoRNP proteins (Parker and Steitz, 1987). It is very likely that all functions of box D in stabilization, nuclear retention and hypermethylation of U3 snoRNA are mediated through its interaction with a specific snoRNP protein(s) (Baserga et al., 1991; Terns et al., 1995). The box C'/D motif of U3 snoRNAs is probably the analog of the box C/D motif of other box C/D snoRNAs. A more precise structure of this RNA element has yet to be elucidated.

The Central Stem is also required for the accumulation of U3 snoRNA and formation of the box C'/D recognition motif (Samarsky and Fournier, 1998). The upper part of the Central Stem of *Cre* U3 snoRNA, similarly to *T. brucei* U3 snoRNA (Hartshorne and Agabian, 1994), is resistant to RNase V1 cleavages in solution. In the WT-2 structure of *Cre* U3 snoRNA, the upper part of the Central Stem is opened and the unfolded 5' strand forms another helix with box B sequences (Fig. 7). Based on its function, the Central Stem of U3 snoRNA might be divided into two distinct domains. While the lower part of the Central Stem is important for shaping the box C'/D protein recognition motif, its upper part likely contributes to the tree-dimensional structure of the Multibranch Loop that functions as an important binding site for snoRNP proteins (Samarsky and Fournier, 1998).

The Multibranch Loop containing the conserved boxes B and C plays an essential role in the function of U3 snoRNA. In the two-dimensional structure of U3 snoRNA, boxes B and C are always close to each other, suggesting that they might function as a single motif similar to the box C'/D motif. In the human U3 snoRNP, the

box C sequences are protected against oligodeoxyribonucleotide-directed RNase H degradation (Parker and Steitz, 1987) and play an essential role in binding the fibrillarin snoRNP protein (Baserga et al., 1991). However, it seems very likely that fibrillarin binding is mediated by snoRNP core proteins that directly bind to the box C region (Lübben et al., 1993). In box C of *Cre* U3 snoRNA, residue C167 was perfectly protected under *in vivo* conditions (Fig. 12). Since this evolutionarily perfectly conserved C residue was completely resistant to *in vivo* modifications in human, *T. brucei*, *S. cereviseae* and *X. laevis* cells as well (Parker and Steitz, 1987; Hartshorne and Agabian, 1994; Jeppesen et al., 1988; Méreau et al., 1997), it might be a good candidate for direct interaction with snoRNP proteins. U3-55k is a protein that has been found to interact with U3 but no other members of the box C/D snoRNA family. Lukowiak et al. have found (Lukowiak et al., 2000) that interaction of the U3-55k protein with U3 RNA *in vivo* is mediated by the conserved box B/C motif. Their *in vitro* RNA binding assays indicated that box C is the primary determinant of this interaction. In Stem/Loop II helices proximal to the Multibranch Loop became sensitive to single strand-specific probes both in the *Cre* (Fig. 12) and yeast (Méreau et al., 1997) U3 snoRNPs, suggesting that these regions may directly or indirectly function in the binding of snoRNP proteins (Fig. 12).

The box B region shows an inconsistent behavior in different organisms. A five nucleotide-long fragment (G109-U113) of box B of *Cre* U3 RNA, most likely due to the formation of a short hairpin, is strongly protected both *in vitro* and *in vivo* (Fig. 7 and 12). Although boxes B of both human and yeast U3 snoRNAs as more sensitive to single strand-specific probes than box C, in human U3 snoRNP the box B sequences proved to be more sensitive to oligodeoxynucleotide-directed RNase H hydrolyses (Parker and Steitz, 1987). Mutation of box B disrupted the interaction of U3-55k protein with U3 RNA to a greater extent *in vivo* than *in vitro* (Lukowiak et al., 2000). The greater importance *in vivo* may reflect the existence of other protein bind box B and by association with U3-55k can stabilize the interaction between U3 RNA and 55k protein. Therefore, it remains unclear whether the box B region contributes to U3 function by taking part in direct interaction(s) with snoRNP proteins or only by facilitating the proper folding of the Multibranch Loop domain of U3 snoRNA.

Recently, it was published that both the box C'/D and the box B/C motif of U3 snoRNAs can be folded into a structure similar to the 5' stem-loop of the U4 snRNA,

which directly binds the yeast Snu13p and its human ortholog 15.5kD proteins (Nottrott et al., 1999; Watkins et al., 2000). In the case of yeast U3 snoRNA it is proved that the RNA binds two 15.5 kD proteins by both motifs (Watkins et al., 2000). In inserts of Fig. 12 we draw the structure of the box C'/D and the box B/C part of *Cre* U3 snoRNA similarly to the structure of the 5' stem-loop of U4 snRNA in order to see how the modification pattern in living cells fits into the proposed structure of a protein binding motif. In both cases increased reactivities of some nucleotides of boxes C and C' in the loop region and partial protection of stem residues were detected. This result suggests the existence of a protein binding structure, but owing to the lack of complete protection at adenines and cytosines we cannot identify exact protein-binding site at the nucleotide level.

**5.4 A pseudoknot in U3 snoRNA?**

The non-conserved stem/loop structures of the 3' domain of U3 snoRNA have been proposed to facilitate RNA folding and processing (Samarsky and Fournier, 1998). In vertebrate, higher and lower plant (*Cre*) U3 snoRNAs, boxes B and C are separated by one, in *S. cerevisiae* U3 snoRNA, by two stem/loop structures (Solymosy and Pollák, 1993). In *Trypanosoma* U3 snoRNA, boxes B and C are located on a large terminal loop (Hartshorne and Agabian, 1994). It is, however, noteworthy that the stem/loop structures separating boxes B and C in yeast U3 snoRNA were found to be dispensable (Fournier et al., 1998). Stem/Loop III of the *Cre* U3 snoRNA, at least in solution, plays an essential role in structure formation of the 3' domain of the RNA (Fig. 10).

During *in vitro* structure probing experiments, the box C region of U3 snoRNAs derived from various organisms was found to be resistant to single strand-specific probes (Méreau et al., 1997; Parker and Steitz, 1987; Jeppesen et al., 1988) and frequently showed RNase V1 cleavages (Méreau et al., 1997; Hartshorne and Agabian, 1994). Interpretation of this unexpected behavior of box C is quite difficult, since it is believed to be single-stranded because they lack a nearby complementarity. However, we noticed that in the WT-1 structure of the *Cre* U3 snoRNA the box C sequences are able to form a five base-pair interaction with the Loop III region, resulting in a pseudoknot structure (Fig. 14). After systematic inspection, we found that, in principle, all known U3 snoRNAs could form a box C - Loop III-type pseudoknot structure (Fig. 14). Several examples demonstrate that RNA pseudoknot structures frequently function

as excellent protein docking signals. The significance of a putative pseudoknot structure of eukaryotic U3 snoRNAs, if any, remains to be established in the future.

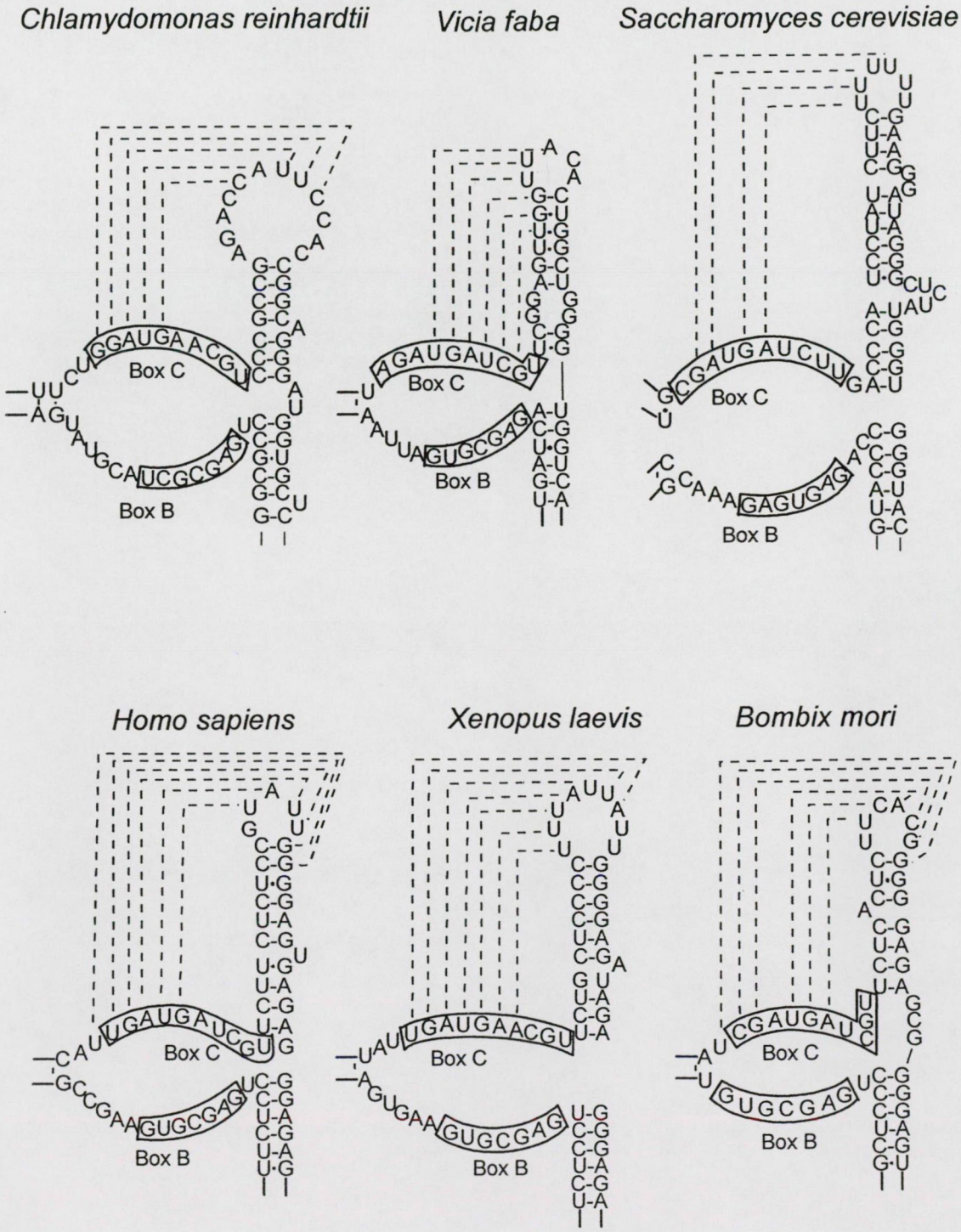


**Figure 14.** Possible pseudoknot formation between C box and Loop III of U3 snoRNAs in different species. Broken lines refer to complementary bases. Conserved nucleotides of boxes B and C are framed.

# 6 REFERENCES


Ajuh, P.M. and Maden, E.B. (1994). Chemical secondary structure probing of two highly methylated regions in Xenopus laevis 28S ribosomal RNA. Biochim. Biophys. Acta *1219*, 89-97.

Antal, M., Kis, M., Mougin, A., Steger, G., Boros, E., Jakab, G., Branlant, C., and Solymosy, F. (2000). Molecular characterization at the RNA and gene levels of U3 snoRNA from a unicellular green alga, *Chlamydomonas reinhardtii*. Nucleic Acids Res. *28*, 2959-68.

Baserga, S.J., Yang, X.D., and Steitz, J.A. (1991). An intact Box C sequence in the U3 snRNA is required for binding of fibrillarin, the protein common to the major family of nucleolar snRNPs. EMBO J. *10*, 2645-2651.

Baserga, S.J., Agentis, T.M., Wormsley, S., Dunbar, D.A., and Lee, S. (1997). Mpp10p, a new protein component of the U3 snoRNP required for processing of 18S rRNA precursors. Nucleic Acids Symp. Ser. 64-67.

Beltrame, M., Henry, Y., and Tollervey, D. (1994). Mutational analysis of an essential binding site for the U3 snoRNA in the 5' external transcribed spacer of yeast pre-rRNA [corrected and republished in Nucleic Acids Res 1994 Nov 25; 22(23):5139-47]. Nucleic Acids Res. *22*, 4057-4065.

Beltrame, M. and Tollervey, D. (1992). Identification and functional analysis of two U3 binding sites on yeast pre-ribosomal RNA. EMBO J. *11*, 1531-1542.

Beltrame, M. and Tollervey, D. (1995). Base pairing between U3 and the pre-ribosomal RNA is required for 18S rRNA synthesis. EMBO J. *14*, 4350-4356.

Borovjagin, A.V. and Gerbi, S.A. (1999). U3 Small Nucleolar RNA is Essential for Cleavage at Sites 1, 2 and 3 in pre-rRNA and Determines which rRNA Processing Pathway is Taken in Xenopus Oocytes. J. Mol. Biol. *286*, 1347-1363.

Borovjagin, A.V. and Gerbi, S.A. (2000). The spacing between functional Cis-elements of U3 snoRNA is critical for rRNA processing. J. Mol. Biol. 2000. Jun. 30;300(1):57-74. *300*, 57-74.

Bortolin, M.L., Ganot, P., and Kiss, T. (1999). Elements essential for accumulation and function of small nucleolar RNAs directing site-specific pseudouridylation of ribosomal RNAs. EMBO J. *18*, 457-469.

Brule, F., Venema, J., Segault, V., Tollervey, D., and Branlant, C. (1996). The yeast Hansenula wingei U3 snoRNA gene contains an intron and its coding sequence co-evolved with the 5' ETS region of the pre-ribosomal RNA. RNA *2*, 183-197.

Busch, H., Reddy, R., Rothblum, L., and Choi, Y.C. (1982). SnRNAs, SnRNPs, and RNA processing. Annu. Rev. Biochem. *51*, 617-654.

Chabot, B. (1994). Synthesis and purification of RNA substrates RNA Processing. Practical Approach Higgins,S.J. Hames,B.D. IRL Press, Oxford, *Volume I*, 1-29.

Chapon, C., Cech, T.R., and Zaug, A.J. (1997). Polyadenylation of telomerase RNA in budding yeast. RNA *3*, 1337-1351.

Ciesiolka, J., Michalowski, D., Wrzesinski, J., Krajewski, J., Krzyzosiak, W.J. (1998). Patterns of cleavages induced by lead ions in defined RNA secondary structure motifs. J. Mol. Biol. *275*, 211-220.

Colley, A., Beggs, J.D., Tollervey, D., and Lafontaine, D.L. (2001). Dhr1p, a putative DEAH-box RNA helicase, is associated with the box C+D snoRNP U3. Mol. Cell Biol. 2000. Oct;20(19):7238-46. *20*, 7238-46.

Craig, N., Kass, S., and Sollner-Webb, B. (1987). Nucleotide sequence determining the first cleavage site in the processing of mouse precursor rRNA. Proc. Natl. Acad. Sci. U. S. A. *84*, 629-633.

Dennis, P.P., Russell, A.G., and Moniz De Sa, M. (1997). Formation of the 5' end pseudoknot in small subunit ribosomal RNA: involvement of U3-like sequences [letter]. RNA *3*, 337-343.

Efstratiadis, A., Vournakis, J.N., Donis-Keller, H., Chaconas, G., Dougall, D.K., and Kafatos, F.C. (1977). End labeling of enzymatically decapped mRNA. Nucleic Acids Res. *4*, 4165-4174.

Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J.P., and Ehresmann, B. (1987). Probing the structure of RNAs in solution. Nucleic Acids Res. *15*, 9109-9128.

Eichler, D.C. and Craig, N. (1994). Processing of eukaryotic ribosomal RNA. Prog. Nucleic Acid. Res. Mol. Biol. *49*, 197-239.

Eschenfeldt,W.H., Puskas,R.S. and Berger,S.L. (1987). Homopolymeric tailing. Methods Enzymol., *152*, 304-305.

Filipowicz, W., Kiss, T., Marshallsay, C., and Waibel, F. (1990). U-snRNA genes, U-snRNAs and U-snRNPs of higher plants. Mol. Biol. Rep. *14*, 125-129.

Fournier, R., Brule, F., Segault, V., Mougin, A., and Branlant, C. (1998). U3 snoRNA genes with and without intron in the Kluyveromyces genus: yeasts can accommodate great variations of the U3 snoRNA 3'-terminal domain. RNA *4*, 285-302.

Ganot, P., Bortolin, M.L., and Kiss, T. (1997a). Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs [see comments]. Cell *89*, 799-809.

Ganot, P., Caizergues-Ferrer, M., and Kiss, T. (1997b). The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. Genes Dev. *11*, 941-956.

Gautier, T., Berges, T., Tollervey, D., and Hurt, E. (1997). Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p and are required for ribosome biogenesis. Mol. Cell Biol. *17*, 7088-98.

Greider, C.W. (1996). Telomere length regulation. Annu. Rev. Biochem. *65:337-65*, 337-365.

Gutell, R.R., Larsen, N., and Woese, C.R. (1994). Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective. Microbiol. Rev. *58*, 10-26.

Hartmuth, K., Raker, V.A., Huber, J., Branlant, C., and Luhrmann, R. (1999). An unusual chemical reactivity of Sm site adenosines strongly correlates with proper assembly of core U snRNP particles. J. Mol. Biol. *285*, 133-147.

Hartshorne, T. (1998). Distinct regions of U3 snoRNA interact at two sites within the 5' external transcribed spacer of pre-rRNAs in Trypanosoma brucei cells. Nucleic Acids Res. *26*, 2541-2553.

Hartshorne, T. and Agabian, N. (1994). A common core structure for U3 small nucleolar RNAs. Nucleic Acids Res. *22*, 3354-3364.

Hartshorne, T. and Toyofuku, W. (1999). Two 5'-ETS regions implicated in interactions with U3 snoRNA are required for small subunit rRNA maturation in Trypanosoma brucei. Nucleic Acids Res. *27*, 3300-3309.

Hughes, J.M., Konings, D.A., and Cesareni, G. (1987). The yeast homologue of U3 snRNA.X. EMBO J. *6*, 2145-2155.

Hughes, J.M. (1996). Functional base-pairing interaction between highly conserved elements of U3 small nucleolar RNA and the small ribosomal subunit RNA. J. Mol. Biol. *259*, 645-654.

Hughes, J.M. and Ares, M., Jr. (1991). Depletion of U3 small nucleolar RNA inhibits cleavage in the 5'X external transcribed spacer of yeast pre-ribosomal RNA and impairsX formation of 18S ribosomal RNA.X. EMBO J. *10*, 4231-4239.

Inoue, T. and Cech, T.R. (1985). Secondary structure of the circular form of the Tetrahymena rRNA intervening sequence: a technique for RNA structure analysis using chemical probes and reverse transcriptase. Proc. Natl. Acad. Sci. U. S. A. *82*, 648-52.

Intine, R.V., Good, L., and Nazar, R.N. (1999). Essential structural features in the schizosaccharomyces pombe pre-rRNA 5' external transcribed spacer [In Process Citation]. J. Mol. Biol. *286*, 695-708.

Jakab, G., Mougin, A., Kis, M., Pollak, T., Antal, M., Branlant, C., and Solymosy, F. (1997). Chlamydomonas U2, U4 and U6 snRNAs. An evolutionary conserved putative third interaction between U4 and U6 snRNAs which has a counterpart in the U4atac-U6atac snRNA duplex. Biochimie, *79*, 387-395.

Jansen, R., Tollervey, D., and Hurt, E.C. (1993). A U3 snoRNP protein with homology to splicing factor PRP4 and G beta domains is required for ribosomal RNA processing. EMBO J. *12*, 2549-2558.

Jeppesen, C., Stebbins-Boaz, B., and Gerbi, S.A. (1988). Nucleotide sequence determination and secondary structure of Xenopus U3 snRNA. Nucleic Acids Res. *16*, 2127-2148.

Kass, S., Tyc, K., Steitz, J.A., and Sollner-Webb, B. (1990). The U3 small nucleolar ribonucleoprotein functions in the first step of preribosomal RNA processing. Cell *60*, 897-908.

Kiss,T., Antal,M. and Solymosy,F. (1987) Plant small nuclear RNAs. II. U6 RNA and a 4.5SI-like RNA are present in plant nuclei. Nucleic Acids Res., *15*, 543-560.

Kiss, T., Marshallsay, C., and Filipowicz, W. (1991). Alteration of the RNA polymerase specificity of U3 snRNA genes during evolution and in vitro. Cell *65*, 517-526.

Kiss, T., and Solymosy, F. (1990). Molecular analysis of a U3 RNA gene locus in tomato: transcription signals, the coding region, expression in transgenic tobacco plants and tandemly repeated pseudogenes. Nucleic Acids Res. *18*, 1941-1949.

Kiss-Laszlo, Z., Henry, Y., Bachellerie, J.P., Caizergues-Ferrer, M., and Kiss, T. (1996). Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. Cell *85*, 1077-1088.

Kiss-Laszlo, Z., Henry, Y., and Kiss, T. (1998). Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. EMBO J. *17*, 797-807.

Kressler, D., Linder, P., and de la Cruz, J. (1999). Protein trans-acting factors involved in ribosome biogenesis in Saccharomyces cerevisiae. Mol. Cell Biol. *19*, 7897-912.

Kunkel, T.A., Roberts, J.D., Zakour, R.A., (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. *154*, 367-82, 367-382.

Lafontaine, D.L. and Tollervey, D. (1999). Nop58p is a common component of the box C+D snoRNPs that is required for snoRNA stability. RNA *5*, 455-467.

Lafontaine, D.L. and Tollervey, D. (2000). Synthesis and assembly of the box C+D small nucleolar RNPs. Mol. Cell Biol. 2000. Apr;20(8):2650-9. *20*, 2650-2659.

Lee, S.J. and Baserga, S.J. (1997). Functional separation of pre-rRNA processing steps revealed by truncation of the U3 small nucleolar ribonucleoprotein component, Mpp10. Proc. Natl. Acad. Sci. U. S. A. *94*, 13536-13541.

Lee, S.J. and Baserga, S.J. (1999). Imp3p and Imp4p, two specific components of the U3 small nucleolar ribonucleoprotein that are essential for pre-18S rRNA processing. Mol. Cell Biol. *19*, 5441-5452.

Liu,M.H., Busch,R.K., Buckley,B. and Reddy,R. (1992) Characterization of antibodies against methyl-pppN cap structure: plant U3 small nucleolar RNA is recognized by these antibodies. Nucleic Acids Res., *20*, 4299-4304.

Lübben, B., Marshallsay, C., Rottmann, N., and Lührmann, R. (1993). Isolation of U3 snoRNP from CHO cells: a novel 55 kDa protein binds to the central part of U3 snoRNA. Nucleic Acids Res. *21*, 5377-5385.

Lück, R., Graf, S., and Steger, G. (1999). ConStruct: a tool for thermodynamic controlled prediction of conserved secondary structure. Nucleic Acids Res. *27*, 4208-4217.

Lukowiak, A.A., Granneman, S., Mattox, S.A., Speckmann, W.A., Jones, K., Pluk, H., Venrooij, W.J., Terns, R.M., and Terns, M.P. (2000). Interaction of the U3-55k protein with U3 snoRNA is mediated by the box B/C motif of U3 and the WD repeats of U3-55k [In Process Citation]. Nucleic Acids Res. 2000. Sep. 15;28(18):3462-71. *28*, 3462-3471.

Lyman, S.K., Gerace, L., and Baserga, S.J. (1999). Human Nop5/Nop58 is a component common to the box C/D small nucleolar ribonucleoproteins. RNA *5*, 1597-604.

Maden, B.E. (1990). The numerous modified nucleotides in eukaryotic ribosomal RNA. Prog. Nucleic Acid. Res. Mol. Biol. *39:241-303.* 241-303.

Marshallsay, C., Kiss, T., and Filipowicz, W. (1990). Amplification of plant U3 and U6 snRNA gene sequences using primers specific for an upstream promoter element and conserved intragenic regions. Nucleic Acids Res. *18*, 3459-3466.

Maxwell, E.S. and Fournier, M.J. (1995). The small nucleolar RNAs. Annu. Rev. Biochem. *64*, 897-934.

Mereau, A., Fournier, R., Gregoire, A., Mougin, A., Fabrizio, P., Luhrmann, R., and Branlant, C. (1997). An in vivo and in vitro structure-function analysis of the Saccharomyces cerevisiae U3A snoRNP: protein-RNA contacts and base-pair interaction with the pre-ribosomal RNA. J. Mol. Biol. *273*, 552-571.

Mougey, E.B., Pape, L.K., and Sollner-Webb, B. (1993). A U3 small nuclear ribonucleoprotein-requiring processing event in the 5' external transcribed spacer of Xenopus precursor rRNA. Mol. Cell Biol. *13*, 5990-5998.

Myslinski, E., Segault, V., and Branlant, C. (1990). An intron in the genes for U3 small nucleolar RNAs of the yeast Saccharomyces cerevisiae. Science *247*, 1213-1216.

Narayanan, A., Lukowiak, A., Jady, B.E., Dragon, F., Kiss, T., Terns, R.M., and Terns, M.P. (1999). Nucleolar localization signals of box H/ACA small nucleolar RNAs. EMBO. J. *18*, 5120-5130.

Newman, D.R., Kuhn, J.F., Shanab, G.M., and Maxwell, E.S. (2001). Box C/D snoRNA-associated proteins: two pairs of evolutionarily ancient proteins and possible links to replication and transcription. RNA 2000. Jun;6(6):861-79. *6*, 861-79.

Nottrott, S., Hartmuth, K., Fabrizio, P., Urlaub, H., Vidovic, I., Ficner, R., and Luhrmann, R. (1999). Functional interaction of a novel 15.5kD [U4/U6.U5] tri-snRNP protein with the 5' stem-loop of U4 snRNA. EMBO. J. *18*, 6119-33.

Parker, K.A. and Steitz, J.A. (1987). Structural analysis of the human U3 ribonucleoprotein particle reveal a conserved sequence available for base pairing with pre-rRNA. Mol. Cell Biol. *7*, 2899-2913.

Pluk, H., Soffner, J., Luhrmann, R., and van Venrooij, W.J. (1998). cDNA cloning and characterization of the human U3 small nucleolar ribonucleoprotein complex-associated 55-kilodalton protein. Mol. Cell Biol. *18*, 488-498.

Porter, G.L., Brennwald, P.J., Holm, K.A., and Wise, J.A. (1988). The sequence of U3 from Schizosaccharomyces pombe suggests structural divergence of this snRNA between metazoans and unicellular eukaryotes [published erratum appears in Nucleic Acids Res 1991 Jun 25;19(12):3484]. Nucleic Acids Res. *16*, 10131-10152.

Raue, H.A. and Planta, R.J. (1995). The pathway to maturity: processing of ribosomal RNA in Saccharomyces cerevisiae. Gene Expr. *5*, 71-77.

Rochaix, J.-D., Mayfield, M., Goldschmidt-Clermont, M. & Erickson, J. (1988). Molecular biology of *Chlamydomonas*, in Plant molecular biology (a practical approach). Shaw, C. H. Ed., IRL Press, Oxford, 253-275.

Samarsky, D.A. and Fournier, M.J. (1998). Functional mapping of the U3 small nucleolar RNA from the yeast Saccharomyces cerevisiae. Mol. Cell Biol. *18*, 3431-3444.

Savino, R. and Gerbi, S.A. (1990). In vivo disruption of Xenopus U3 snRNA affects ribosomal RNA processing. EMBO J. *9*, 2299-2308.

Schimmang, T., Tollervey, D., Kern, H., Frank, R., and Hurt, E.C. (1989). A yeast nucleolar protein related to mammalian fibrillarin is associated with small nucleolar RNA and is essential for viability. EMBO J. *8*, 4015-4024.

Segault, V., Mougin, A., Gregoire, A., Banroques, J., and Branlant, C. (1992). An experimental study of Saccharomyces cerevisiae U3 snRNA conformation in solution. Nucleic Acids Res. *20*, 3443-3451.

Sharma, K. and Tollervey, D. (1999). Base pairing between U3 small nucleolar RNA and the 5' end of 18S rRNA is required for pre-rRNA processing. Mol. Cell Biol. *19*, 6012-6019.

Shimba, S., Buckley, B., Reddy, R., Kiss, T., and Filipowicz, W. (1992). Cap structure of U3 small nucleolar RNA in animal and plant cells is different. gamma-Monomethyl phosphate cap structure in plant RNA. J. Biol. Chem. *267*, 13772-13777.

Solymosy, F. and Pollak, T. (1993). Uridylate-Rich Small Nuclear RNAs (UsnRNAs), Their Genes ans Pseudogenes, and UsnRNPs in Plants: Structure and Function. A Comparative Approach. Critical Review in Plant Sciences *12*, 275-369.(Abstract)

Speckmann, W., Narayanan, A., Terns, R., and Terns, M.P. (1999). Nuclear retention elements of U3 small nucleolar RNA. Mol. Cell Biol. *19*, 8412-8421.

Staley, J.P. and Guthrie, C. (1998). Mechanical devices of the spliceosome: motors, clocks, springs, and things. Cell *92*, 315-26.

Terns, M.P., Grimm, C., Lund, E., and Dahlberg, J.E. (1995). A common maturation pathway for small nucleolar RNAs. EMBO J. *14*, 4860-4871.

Terns, M.P. and Dahlberg, J.E. (1994). Retention and 5' cap trimethylation of U3 snRNA in the nucleus. Science *264*, 959-961.

Tollervey, D., Lehtonen, H., Carmo-Fonseca, M., and Hurt, E.C. (1991). The small nucleolar RNP protein NOP1 (fibrillarin) is required for pre- rRNA processing in yeast. EMBO J. *10*, 573-583.

Tyc, K. and Steitz, J.A. (1989). U3, U8 and U13 comprise a new class of mammalian snRNPs localized in the cell nucleolus. EMBO J. *8*, 3113-3119.

Venema, J., and Tollervey, D. (1999). Ribosome synthesis in Saccharomyces cerevisiae. Annu. Rev. Genet. *33:261-311*, 261-311.

Venema, J., Vos, H.R., Faber, A.W., van Venrooij, W.J., and Raue, H.A. (2001). Yeast Rrp9p is an evolutionarily conserved U3 snoRNP protein essential for early pre-rRNA processing cleavages and requires box C for its association. RNA 2000. Nov;6(11):1660-71. *6*, 1660-71.

Watkins, N.J., Gottschalk, A., Neubauer, G., Kastner, B., Fabrizio, P., Mann, M., and Lührmann, R. (1998a). Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA- binding protein, are present together with Garlp in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure. RNA *4*, 1549-1568.

Watkins, N.J., Newman, D.R., Kuhn, J.F., and Maxwell, E.S. (1998b). In vitro assembly of the mouse U14 snoRNP core complex and identification of a 65-kDa box C/D-binding protein. RNA *4*, 582-593.

Watkins, N.J., Segault, V., Charpentier, B., Nottrott, S., Fabrizio, P., Bachi, A., Wilm, M., Rosbash, M., Branlant, C., and Luhrmann, R. (2000). A common core RNP structure shared between the small nucleoar box C/D RNPs and the spliceosomal U4 snRNP [In Process Citation]. Cell 2000. Oct. 27;103(3):457-66. *103*, 457-466.

Wiederkehr, T., Pretot, R.F., and Minvielle-Sebastia, L. (1998). Synthetic lethal interactions with conditional poly(A) polymerase alleles identify LCP5, a gene involved in 18S rRNA maturation. RNA *4*, 1357-72.

Wise, J.A. and Weiner, A.M. (1980). Dictyostelium small nuclear RNA D2 is homologous to rat nucleolar RNA U3 and is encoded by a dispersed multigene family. Cell *22*, 109-118.

Wormsley, S., Samarsky, D.A., Fournier, M.J., and Baserga, S.J. (2001). An unexpected, conserved element of the U3 snoRNA is required for Mpp10p association. RNA 2001. Jun;7(6):904-19. *7*, 904-19.

Xia, L., Watkins, N.J., and Maxwell, E.S. (1997). Identification of specific nucleotide sequences and structural elements required for intronic U14 snoRNA processing. RNA *3*, 17-26.

Zaug, A.J. and Cech, T.R. (1995). Analysis of the structure of Tetrahymena nuclear RNAs in vivo: telomerase RNA, the self-splicing rRNA intron, and U2 snRNA. RNA *1*, 363-374.

## Acknowledgements

I am highly indebted to Prof. Ferenc Solymosy for providing me with the possibility of starting my scientific carrier in his group. His excellent scientific guidance, and constructive criticism helped me from the first moment until now.

I have much to be thankful to Prof Tamás Kiss. He introduced me into the most up-to-date laboratory methods and into the most efficient way of experimentation. I appreciated very much being member of his research team.

I thank the referees Prof. András Lipták and Prof. Ernő Duda for the critical reading of my thesis.

I thank Prof. Christiane Branlant and her laboratory for her fruitful cooperation in structure probing experiments and for the publication of the results.

I thank my colleagues in Szeged, Mihály Kis, Tamás Pollák, Éva Boros, Beáta Erika Jády, Arnold Mihály Kis, and Ildikó Borka for creating a pleasant atmosphere and for their cooperation in my laboratory work.

I am grateful to my colleagues in Nancy, Annie Mougin, Ágnes Méreau and Klaus Hartmuth, who gave me a hand in managing everyday life and laboratory work during my stay in France.

I would like to thank my parents and sons for their love and tolerance during my scientific work.